Imperial College London
Department of Computing

Sequent Calculi with Context Restrictions and Applications to Conditional Logic

Björn Lellmann

October 28, 2013

Supervised by Dirk Pattinson

Submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy in Computing of Imperial College London and the Diploma of Imperial College London

Declaration

I hereby declare that all work reported in this thesis is my own unless stated otherwise.

Björn Lellmann

Abstract

In this thesis we consider generic tools and techniques for the proof-theoretic investigation of not necessarily normal modal logics based on minimal, intuitionistic or classical propositional logic. The underlying framework is that of ordinary symmetric or asymmetric two-sided sequent calculi without additional structural connectives, and the point of interest are the logical rules in such a system. We introduce the format of a sequent rule with context restrictions and the slightly weaker format of a shallow rule. The format of a rule with context restrictions is expressive enough to capture most normal modal logics in the S5 cube, standard systems for minimal, intuitionistic and classical propositional logic and a wide variety of non-normal modal logics. For systems given by such rules we provide sufficient criteria for cut elimination and decidability together with generic complexity results. We also explore the expressivity of such systems with the cut rule in terms of axioms in a Hilbert-style system by exhibiting a corresponding syntactically defined class of axioms along with automatic translations between axioms and rules. This enables us to show a number of limitative results concerning amongst others the modal logic S5. As a step towards a generic construction of cut free and tractable sequent calculi we then introduce the notion of cut trees as representations of rules constructed by absorbing cuts. With certain limitations this allows the automatic construction of a cut free and tractable sequent system from a finite number of rules. For cases where such a system is to be constructed by hand we introduce a graphical representation of rules with context restrictions which simplifies this process. Finally, we apply the developed tools and techniques and construct new cut free sequent systems for a number of Lewis' conditional logics extending the logic V. The systems yield purely syntactic decision procedures of optimal complexity and proofs of the Craig interpolation property for the logics at hand.

To my family.

Acknowledgements

First and foremost I would like to thank my supervisor, Dirk Pattinson, for the continuous support, both moral and technical, during the whole of my PhD, for the long fruitful discussions and for giving me the chance to embark on this adventure in the first place. I greatly profited from his comments on a draft of this thesis and his commitment to his students in general. Dirk also made possible my visit to Australia, which significantly contributed to the results presented in this thesis. Without his encouragement, confidence and trust none of this would have happened.

I am grateful to Marek Sergot who took over the role of the official supervisor and provided much appreciated support during the final stage of my PhD.

Also I would like to express my gratitude to the members of the research group at Imperial, who apart from providing a very stimulating research environment also guaranteed ample relaxation (not only) after the seminars: Faris Abu-Saleh, Nick Bezhanishvili, Clemens Kupke, Rob Myers, and in particular Fredrik Dahlqvist and Florian Widman, who being also my officemates had to endure my bouncing my ideas of them increasingly often.

Many other people whom I met at seminars or conferences contributed to this thesis in a direct or indirect way by providing helpful comments, suggestions or other intellectual stimulation, including Régis Alenda, Michael Arndt, James Brotherston, Agata Ciabattoni, Rajeev Goré, Ori Lahav, Tadeusz Litak, George Metcalfe, Nicola Olivetti, Dmitry Tishkovsky, Alwen Tiu and Anna Zamansky.

The Australian National University generously hosted my month-long stay in Canberra, during which most of the limitative results took form.

I gratefully acknowledge the generous financial support provided by the EPSRC-project EP/H016317/1, without which this work would not have been possible.

I would also like to thank my parents for their continuous support, encouragement and reassurance, and my brother for setting the example he did.

Last but most certainly not least, I cannot imagine what I would have done without the all-encompassing support of my wonderful partner Franziska Siess. Thank you for accepting the challenge to move to strange places and for providing a home so hard to leave in the morning!

Contents

List of Tables

List of Figures

1 Introduction

1.1 The Questions

The emergence of ever new propositional modal logics in particular in the field of Computer Science calls for the development of generic methods to deal with such logics in a systematic way. Not only can such generic methods provide a starting point for more specific treatments of particular logics, they can also serve to reveal general principles which in turn may guide the development and discovery of new logics with good properties. Unfortunately, the matter is slightly complicated by the fact that the introduced logics often are based on a semantics different from the standard Kripke semantics and hence often make use of non-normal modalities. Moreover, e.g. in the context of logics describing aspects of computation, there is an increasing number of modal logics based on non-classical propositional logic, and in particular on intuitionistic propositional logic (see e.g. [Wij90, FM97, GGGP11]). To be maximally useful a generic treatment should therefore try to incorporate non-normal modalities and non-classical propositional base logics as well.

One of the main questions when dealing with such logics, and therefore a good candidate for a generic treatment, is that of decidability and complexity of the logic in question. Proof theory, and in particular its branch dedicated to the study of sequent calculi, has turned out to be a valuable tool for tackling these questions. This is mainly due to the fact that cut free sequent calculi, the main objects of study in this field, usually enjoy the subformula property which in turn often can be exploited to derive decidability and complexity results. Combining all the mentioned aspects it is therefore of increasing interest to study generic methods in the proof theory, and in particular the theory of sequent calculi, for such logics. This gives rise to the first main question we are going to consider, a question of distinctively *constructive* character:

Question 1.1.1. Are there generic methods to automatically generate 'nice' sequent calculi for modal logics?

Ideally, such generic methods would yield sequent calculi which can be used in generic decision procedures for the logics in question. But of course there are (even reasonably simple) undecidable modal logics [KNSS95], so we should not expect to find methods which work for arbitrary modal logics. In order to delineate the domain where the envisaged generic methods have a chance of succeeding it is therefore imperative to investigate limiting factors as well.

Thus immediately connected with our first question is the second main question considered here, the character of which is best described as *limitative*:

Question 1.1.2. Exactly which modal logics have 'nice' sequent calculi?

Obviously, the answers to these questions depend on at least two major factors. On the one hand we need to know which format the modal logics originally are given in. On the other hand it is necessary to clarify what exactly we mean by a 'nice' sequent calculus. While for normal modal logics the semantics offers a convenient starting point (see e.g. [Neg05, Lah13]), due to the plethora of different semantics for non-normal modal logics or logics based on non-classical propositional logic this seems not to be the optimal approach in our setting. On the other hand, the purely syntactic formulation of modal logics in terms of axioms for a Hilbert-style system can be and very often is used to give a relatively concise and intuitive presentation of the logic under scrutiny. Moreover, this formulation is independent of the semantics and can be used also if the underlying propositional logic is non-classical. For these reasons here we take the formulation in terms of axioms as the point of departure.

Regarding the second major factor, the question what exactly is meant by a 'nice' sequent calculus, things become a little bit more complicated. The first issue is that there are many different formats of 'sequent calculi' for modal logics. Apart from extensions of the original sequent calculi for propositional logic introduced by Gentzen [Gen34] with new logical rules for the modalities [Gob74, Lei81] there are the formats of *labelled sequent* calculi [Neg05], *hypersequent* calculi [Avr96, Pog08], *nested sequent* calculi [Brü09, Pog09] or *display logic* [Bel82, Kra96] just to name a few. While some of these extensions are very powerful, it should be noted that this comes at the price of considerable additional machinery, typically mirrored by less efficient decision procedures as well. Of course ultimately we would like to have generic treatments incorporating all of the mentioned frameworks, but this should perhaps be seen more as a research programme, founded on detailed studies for each framework on its own. As a starting point for such a programme it seems wise to first consider frameworks with little additional machinery and to try to fully understand and exploit the power of one formal framework before considering extensions or different frameworks. For this reason here we are going to stick to the first mentioned framework of standard sequent calculi with new logical rules for the modalities, not because the other extensions are not interesting or fruitful – they most certainly are – but because this extension is one of the simplest and requires the least additional machinery.

Having settled the question of what kind of sequent calculus we are going to consider we need to clarify which properties a 'nice' sequent calculus should possess. There is a lot of philosophically minded debate about this question (see e.g. [Wan02, Pog11]), but since here we are mainly concerned with the computational properties of the sequent calculi we are not going to enter this discussion. But even ignoring all philosophical considerations, the fact that we are aiming for decidability of the calculi in question gives some clues about which properties we should demand of such a system. Since the subformula property is one of the standard tools for proving decidability results and since the presence of the cut rule tends to destroy this property, surely at least admissibility of the cut rule should be one of the criteria for a 'nice' system. On the other hand, this can not be the only such condition, since given a Hilbert-style system we may simply introduce zero-premiss sequent rules $\overline{\Rightarrow A}$ for every theorem A of the Hilbert-style system. The cut rule is trivially admissible in the resulting sequent system, but obviously the computational properties of this system are no better than those of the original Hilbert-system, so we have not gained anything. In order to prevent trivial systems as the one mentioned, we will restrict the format of the sequent rules in a suitable way. Ideally, the rule format should strike a sensible balance between tractability and expressivity: it should allow for a generic proof theoretic treatment, ideally resulting in relatively efficient decision procedures, while at the same time being general enough to cover a wide variety of logics. Here we introduce and propose the format of a *rule with context restrictions*, an abstraction of the standard sequent rules for modal logics such as K, KT or $\mathsf{S4}$. As such this format is general enough to capture many of the standard normal modal logics, all logics axiomatised by non-iterative modal axioms over classical propositional logic and many non-normal logics based on intuitionistic or classical propositional logic. At the same time we can give reasonably simple criteria on the rule set ensuring that generic syntactic cut elimination or complexity theorems hold. Since ultimately we would like to generate cut free systems we will take these criteria as a guide as to what constitutes a 'nice' sequent system. This finally allows us to make Question 1.1.1 more precise:

Question 1.1.3. Are there generic methods to automatically generate sets of rules with context restrictions satisfying the generic criteria for cut elimination and complexity from sets of axioms for a Hilbert-style system?

Regarding our second question we will need to take into account that unfortunately our generic criteria for cut elimination are only sufficient and not necessary for admissibility of the cut rule. While due to the fact that syntactic cut elimination is only one method of proving cut admissibility this situation might be expected, it also means that we cannot use these criteria to characterise the class of (sets of) axioms corresponding to (sets of) rules with restrictions which generate cut free systems. But even so we may still try to limit this class from above using the fact that adding the cut rule to a sequent calculus does not jeopardise soundness. This gives us the following more precise formulation of Question 1.1.2:

Question 1.1.4. Exactly which sets of axioms for a Hilbert-style system correspond to sets of rules with context restrictions?

We will see that this question can be answered by defining a purely syntactically characterised class of axioms for Hilbert-style systems and translations from axioms in this class into equivalent rules with restrictions and vice versa. This also allows us to begin to extend a research programme originally developed in the field of substructural logics to the realm of modal logics, namely the programme of characterising logics according to the necessary proof-theoretic strength of corresponding sequent-style derivation systems (see e.g. [CGT08, CGT12]). In this spirit we will see limitative results about a number of standard normal modal logics including the logics $\mathsf{KT}, \mathsf{K4}, \mathsf{B}, \mathsf{GL}$ and $\mathsf{S5}$. Even though these results should only be considered a first step towards a comprehensive characterisation, they might also provide a small step towards a formal explanation of why certain logics (most notably $\mathsf{S5}$) seem to require additional machinery on top of the format of 'standard' sequent calculi, machinery such as labelled sequents, hypersequents or nested sequents.

While Question 1.1.4 thus has a reasonably comprehensive answer, the situation for Question 1.1.3 is somewhat less satisfying. While we will see a number of generic results, we will not see a *fully automatic* procedure to turn a set of axiom into a cut free sequent system given by rules of our format which also gives rise to efficient decision procedures. In many cases we will still need to manually check that the rule sets satisfy certain criteria. The fact that it is not always easy to check whether these criteria are satisfied and that for some sets of axioms this semi-automatic procedure does not yield a satisfactory set of sequent rules furthermore suggests the development of tools to aid the manual construction of cut-free sequent calculi from sets of axioms. With this in mind we will develop a graphical tool for manually manipulating sequent rules and absorbing cuts into the rule set. Finally, we will put the introduced tools and techniques to work and construct new cut-free sequent calculi for a number of strong systems of conditional logic.

1.2 Contribution and Organisation of the Thesis

In summary, the main contributions reported in this thesis are the following.

1. The conceptual development of the notions of shallow rules and rules with context restrictions (Definition 2.3.3) together with generic cut elimination and complexity theorems for sequent calculi given by such rules (Theorems 2.4.16, 2.7.5, 2.7.8) which significantly extend the results from [SP09, PS09].

2. The identification and purely syntactical characterisation of classes of axioms for a Hilbert-style system corresponding to shallow rules or rules with restrictions including translations from axiom into rules and vice versa (Theorem 3.3.18).

3. The application of these characterisations to derive limitative results for a number of modal logics including GL (Theorem 3.4.16) and $\mathsf{S5}$ (Theorem 3.4.18).

4. The introduction of the notion of cut trees (Definition 4.1.3) to automate the absorption of cuts into a set of rules which under certain conditions gives an automatic construction

of a cut-free sequent calculus suitable for deciding the logic (Theorem 4.1.19 and Corollaries 4.1.20 and 4.1.21).

5. The introduction of a graphical tool for the manipulation of sequent rules (Definition 4.2.14) which aids the process of manually absorbing cuts into a set of rules.

6. New cut free sequent calculi for a number of conditional logics extending Lewis' logic $\mathbb{V}$ formulated using the comparative plausibility operator or the strong counterfactual implication (Definitions 5.2.1, 5.3.3, 5.4.4) which yield syntactic decision procedures of optimal complexity and facilitate proofs of the Craig interpolation property (Theorem 5.5.4 and Corollary 5.5.7).

The remainder of the work is organised as follows.

After a brief recapitulation of some basic notions and notations in Section 1.4 we delve into the main part of the thesis in **Chapter 2**. Here we introduce the fundamental notions including the rule formats of one-step rules, shallow rules and rules with context restrictions in Section 2.3. This is followed by a close investigation of the sequent calculi generated by sets of such rules, including the criteria for generic cut elimination in Sections 2.4 and 2.5, admissibility of the contraction rule in a slightly modified rule set in Section 2.6 and decidability and complexity issues in Section 2.7.

Chapter 3 is dedicated to the study of the connection between rules with context restrictions and axioms for Hilbert-style systems. Following some preparatory considerations in Section 3.1 we define the syntactic format of a *translatable clause* in Section 3.2 and provide a translation from axioms of this format into sequent rules with context restrictions. The translation from sequent rules with context restrictions into sets of translatable clauses for monotone modalities is given in Section 3.3. Together, these two translations yield characterisations of the classes of axioms corresponding to each of the formats of one-step rules, shallow rules and rules with context restrictions. These characterisations are summarised in Table 3.2 on page 96. In the following Section 3.4 we put the characterisations to work and prove limitative results about a number of standard normal modal logics.

Generic methods and tools for the construction of cut-free sequent systems from a finite set of rules with context restrictions are the subject of **Chaper 4**. In Section 4.1 we consider the method of absorbing (principal) cuts into the rule set and identify a representation of the resulting rules in terms of *cut trees*. Provided the corresponding rule sets satisfy certain criteria these representations can be used in the generic decision procedures introduced earlier and thus provide a step towards generic decidability and complexity results. For cases where these criteria are not satisfied we then introduce in Section 4.2 a graphical representation of sequents and sequent rules by *doodles* resp. *rule doodles*, which can be used in the manual construction of such a cut absorbing rule set. Examples considered in this chapter include Elgesem's logic of agency and ability and weak systems of conditional logic.

Finally, in **Chapter 5** we apply these methods in the construction of new cut free sequent calculi for strong systems of conditional logic formulated using the comparative plausibility operator. After a brief introduction to conditional logics with sphere semantics in Section 5.1 the basic systems are constructed using the representation of sequent rules as rule doodles in Section 5.2 for the basic logic $\mathbb{V}_{\preccurlyeq}$ and in Section 5.3 for extensions of this logic. These calculi are then adapted to the strong counterfactual implication as primitive connective in Section 5.4 and used to prove new interpolation results in Section 5.5.

Each chapter closes with a short section titled 'Notes' containing further discussions about related approaches, literature or open problems. Pointers to the relevant notions and definitions can be found in the index.

1.3 Publications

This work is based on the following publications.

1. Björn Lellmann and Dirk Pattinson. Cut elimination for shallow modal logics. In Kai Brünnler and George Metcalfe, editors, *TABLEAUX 2011*, volume 6793 of *LNAI*, pages 211–225. Springer-Verlag Berlin Heidelberg, 2011.

2. Björn Lellmann and Dirk Pattinson. Graphical construction of cut-free sequent systems suitable for backwards proof search (extended abstract). In Renate A. Schmidt and Fabio Papacchini, editors, *Proceedings of the 19th Automated Reasoning Workshop*, pages 17–18, Manchester, 2012. School of Computer Science, The University of Manchester.

3. Björn Lellmann and Dirk Pattinson. Sequent systems for Lewis' conditional logics. In Luis Fariñas del Cerro, Andreas Herzig, and Jerome Mengin, editors, *JELIA 2012*, volume 7519 of *LNCS*, pages 320–332. Springer-Verlag Berlin Heidelberg, 2012.

4. Björn Lellmann and Dirk Pattinson. Constructing cut free sequent systems with context restrictions based on classical or intuitionistic logic. In Kamal Lodaya, editor, *ICLA 2013*, volume 7750 of *LNAI*, pages 148–160. Springer-Verlag Berlin Heidelberg, 2013.

5. Björn Lellmann and Dirk Pattinson. Correspondence between modal Hilbert axioms and sequent rules with an application to S5. In Didier Galmiche and Dominique Larchey-Wendling, editors, *TABLEAUX 2013*, volume 8123 of *LNCS*, pages 219–233. Springer-Verlag Berlin Heidelberg, 2013.

1.4 Notation and Other Preliminaries

Notation. We use the following standard notation.

$$\begin{aligned}
\mathbb{N} &= \{0, 1, 2, 3, \dots\} \\
[n] &= \{1, 2, \dots, n\} \text{ for } n \in \mathbb{N} \text{ with } n > 0 \\
[0] &= \emptyset \\
\lfloor r \rfloor &= \max\{n \in \mathbb{N} \mid n \leq r\} \text{ for non-negative rational } r \\
\lceil r \rceil &= \min\{n \in \mathbb{N} \mid n \geq r\} \text{ for non-negative rational } r \\
|A| &= \text{cardinality of the set } A \\
\mathfrak{P}(A) &= \text{powerset of the set } A
\end{aligned}$$

Complexity Theory. While not fundamental to the understanding of this work, we make use of some basic notions from computational complexity theory for which we briefly recall the main intuitions. The complexity classes occurring are based on the notion of a deterministic or non-deterministic *Turing machine*, where a *deterministic* Turing machine intuitively on receiving a string over a finite alphabet as input calculates simple functions such as boolean addition and multiplication on the cells of a *work tape* according to a fixed set of instructions and depending on the input. A *computation* of such a machine on a given input is *successful* if it eventually stops and returns 'yes'. A *non-deterministic* Turing machines in addition can make existential or universal *guesses* by non-deterministically writing a symbol in a cell on its work tape. A computation of a non-deterministic Turing machine beginning with an *existential* guess is successful if at least one of the immediate subcomputations is successful, whereas for a computation beginning with a *universal* guess to be successful all immediate subcomputations need to be successful. The *time* and *space* needed by the computation are measured in terms of the size n of (the encoding of) the input and are given by the maximal number of steps a computation executes on a given input resp. the amount of cells the computation uses on the work tape. Time and space required to solve a problem are called *polynomial* if they are bounded by a polynomial, i.e. by n^k for some $k \in \mathbb{N}$ and *exponential* if they are bounded by an exponential function, i.e. by 2^{n^k} for some $k \in \mathbb{N}$. The considered complexity classes are given in Table 1.1. The relations between these classes are

$$\text{coNP} \subseteq \text{APTIME} = \text{PSPACE} \subseteq \text{EXPTIME} \subseteq \text{NEXPTIME}\,.$$

Detailed treatments of these notions can be found e.g. in [Pap94, AB09].

coNP	the class of decision problems accepted in *polynomial time* by a *non-deterministic* Turing machine making *universal* guesses
APtime	the class of decision problems accepted in *polynomial time* by a *non-deterministic* Turing machine making *universal and existential* guesses
Pspace	the class of decision problems accepted in *polynomial space* by a *deterministic* Turing machine
Exptime	the class of decision problems accepted in *exponential time* by a *deterministic* Turing machine
NExptime	the class of decision problems accepted in *exponential time* by a *non-deterministic* Turing machine making *existential* guesses

Table 1.1: Overview over the considered complexity classes.

2 Sequent Systems and Cut Elimination

We begin our investigations by defining three different formats of sequent rules, those of one-step rules, shallow rules and rules with context restrictions. While the general format of one-step rules has been investigated before in the context of rank-1 modal logics [PS08, PS10] and the format of shallow rules is a natural extension of this format, which captures e.g. the standard propositional rules, the most general of these formats, that of rules with context restrictions, is motivated by standard sequent calculi for common normal modal logics. In such calculi one often restricts the context formulae in the premisses to formulae of a certain format, e.g. to boxed formulae. Furthermore, usually only one layer of modalities is introduced in the principal formulae of the rule. These two properties will be the main ingredients of the notion of a rule with context restrictions.

After a brief introduction of the fundamental notions in Section 2.1 we will informally consider some intuitions in Section 2.2 using standard sequent rules for some well-known modal logics. This will be followed by a formal introduction of the rule format in Section 2.3 before we take a look at one of the main issues when investigating sequent calculi: cut elimination. Standard syntactical proofs of cut elimination by stepwise transformation of derivations usually involve two or more nested inductions and a plethora of cases. Moreover, adapting the standard proofs to a new calculus often involves ad hoc constructions or modifications to ensure that all the cases go through. These two issues might account for the fact that syntactic cut elimination proofs are notoriously error-prone, as witnessed e.g. by the quest for a correct cut elimination proof for the sequent calculus for Gödel Löb logic GL [Lei81, Moe01, Val83, GR08]. Thus, similar in spirit to the methods used in display logic [Bel82] and building on the characterisations in [PS08, PS10] we try to simplify this process in Sections 2.4 and 2.5 by identifying reasonably simple syntactic criteria on rule sets, which are verifiable on a rule-by-rule basis, and which are sufficient to allow a generic proof of cut elimination to go through.

If we are interested in not only constructing a cut-free sequent calculus, but also in using it to decide which formulae are valid in the corresponding logic, we normally also would like to show that the structural rule of Contraction is admissible. In Section 2.6 we will see how cut-free systems given by sets of rules with context restrictions can be slightly modified in the spirit of the $\mathsf{G3}$-systems of [Kle52] to ensure that this is the case. Since the resulting cut-free sequent calculi enjoy the subformula property, they often can be used to decide which formulae are derivable. In Section 2.7 we give generic decision procedures for such sequent calculi along with generic complexity results.

The format of rules with restrictions together with a variant of the cut elimination theorem, admissibility of contraction and decidability has been published in [LP13a]. Analogous results for shallow rules were published in [LP11].

2.1 Connectives, Formulae and Sequents

Let us first consider the basic notions. Although the emphasis of this work lies on sequent systems for *modal* logics, from a syntactical point of view the modalities are on par with the standard boolean connectives. Thus we will take a slightly more general perspective and consider modalities as *connectives*.

Definition 2.1.1. Let Λ be a finite set of *connectives*, that is symbols with associated *arities* in $\mathbb{N}$. Furthermore let Var be a countable set of propositional variable symbols. The set $\mathcal{F}(\Lambda)$ of *formulae over* Λ is defined by $\mathcal{F}(\Lambda) \ni \varphi ::= p \mid \heartsuit(\varphi_1, \dots, \varphi_n)$ where $p \in \mathsf{Var}$ and $\heartsuit \in \Lambda$ has arity n. If F is a set of formulae we write $\Lambda(F)$ for the set $\{\heartsuit(A_1, \dots, A_n) \mid \heartsuit \in \Lambda\ n\text{-ary and } A_i \in F \text{ for } i \in \{1, \dots, n\}\}$. For a formula A we write $\mathsf{var}\,(A)$ for the set of propositional variable symbols occurring in A. We sometimes abbreviate finite sequences $A_1, \dots, A_n$ of formulae to $\vec{A}$.

Usually we stipulate that the binary *boolean connectives* $\to, \wedge, \vee$ and the 0-ary connective $\bot$ are in Λ, and we write the former ones in infix notation. We use the standard definitions

$$\begin{aligned} \top &:= \bot \to \bot \\ \neg A &:= A \to \bot \\ A \leftrightarrow B &:= (A \to B) \wedge (B \to A)\ . \end{aligned}$$

Furthermore we adopt the standard conventions concerning the binding strength of the unary and binary boolean connectives to economise on brackets with the order (from strongest to weakest binding strength): $\neg, \wedge, \vee, \to, \leftrightarrow$. Connectives other than the boolean connectives are also called *modal connectives* or simply *modalities*. The notions of rank and modal nesting depth are defined as usual and give rise to the important notions of shallow and rank-1 formulae.

Definition 2.1.2. For a set Λ of connectives the *rank* of a formula A over Λ is the number $\mathsf{rk}\,(A)$ defined by $\mathsf{rk}\,(p) := 0$ for every propositional variable p and $\mathsf{rk}\,(\heartsuit(A_1, \dots, A_n)) := \max(\{0\} \cup \{\mathsf{rk}\,(A_i) \mid 1 \leq i \leq n\}) + 1$ for $n \geq 0$ and $\heartsuit$ an n-ary connective. The *modal nesting* *depth* or *modal rank* of formulae in $\mathcal{F}(\Lambda)$ is defined by $\mathsf{mrk}\,(p) := \mathsf{mrk}\,(\bot) := 0$ and

$$\mathsf{mrk}\,(\heartsuit(A_1,\dots,A_n)) := \begin{cases} \max\{\mathsf{mrk}\,(A_i) \mid 1 \le i \le n\} & \text{if } \heartsuit \in \{\to,\wedge,\vee\} \\ \max(\{0\} \cup \{\mathsf{mrk}\,(A_i) \mid 1 \le i \le n\}) + 1 & \text{otherwise.} \end{cases}$$

A formula A is called *non-iterative* if $\mathsf{mrk}\,(A) \le 1$. A propositional variable occurs on the *top level* of a formula if it occurs at least once in the formula and none of its occurrences in the formula are in the scope of a modality. A formula A is a *rank-1* formula if $\mathsf{mrk}\,(A) = 1$ and no propositional variable occurs on its top level.

Thus the rank of a formula is the maximal nesting depth of connectives including the boolean connectives, whereas the modal rank only counts the modalities. In particular, rank-1 formulae are non-iterative formulae as well.

Example 2.1.3. We consider formulae over the set $\Lambda_\Box := \{\to,\wedge,\vee,\bot,\Box\}$ of connectives.

1. The formula $\Box p \wedge \Box q \to \Box(p \wedge q)$ has rank 3 and modal rank 1. Thus it is a non-iterative formula. Since moreover all occurring propositional variables are in the scope of a modality it is also a rank-1 formula.

2. The formula $(\mathsf{T}) = \Box p \to p$ has rank 2 and modal rank 1 and thus is a non-iterative formula. Since the variable p occurs on the top level of (T) it is not a rank-1 formula.

3. The formula $(\mathsf{4}) = \Box\Box p \to \Box p$ has rank 3 and modal rank 2. Thus it is neither a non-iterative nor a rank-1 formula.

We will also make use of the standard notion of a substitution:

Definition 2.1.4. For a finite set Λ of connectives a Λ-*substitution* is a function $\sigma : \mathsf{Var} \to \mathcal{F}(\Lambda)$. If the set Λ is clear from the context we also simply say that such a σ is a *substitution.* As usual we extend substitutions to functions on $\mathcal{F}(\Lambda)$ by the clauses

$$\sigma(\heartsuit(A_1,\dots,A_n)) := \heartsuit(\sigma(A_1),\dots,\sigma(A_n))$$

for every $n > 0$ and n-ary connective $\heartsuit$ from Λ and also write $A\sigma$ for $\sigma(A)$.

We are mainly going to consider modal logics based on some form of propositional logic. This motivates the following definition of a logic as a set of formulae including the theorems of the respective propositional logic.

Definition 2.1.5. Let Λ be a finite set of connectives including the boolean connectives. A Λ-*logic based on classical (resp. intuitionistic resp. minimal) propositional logic* is a set $\mathcal{L}$ of formulae from $\mathcal{F}(\Lambda)$ which

1. includes all theorems of classical (resp. intuitionistic resp. minimal) propositional logic
2. is closed under substitution: if $A \in \mathcal{L}$ and σ is a substitution, then $A\sigma \in \mathcal{L}$
3. is closed under *modus ponens*: if $A \to B \in \mathcal{L}$ and $A \in \mathcal{L}$, then $B \in \mathcal{L}$.

If the set Λ of connectives and the underlying propositional logic are clear from the context, we also simply speak of a *logic*. Keeping to the standard notations we also write $\models_{\mathcal{L}} A$ for $A \in \mathcal{L}$ and $\not\models_{\mathcal{L}} A$ for $A \notin \mathcal{L}$ and say that A is *$\mathcal{L}$-valid* if $\models_{\mathcal{L}} A$. If $\Lambda = \{\Box\}$ we say that a Λ-logic is *normal* if it contains the formulae $\Box p \wedge \Box q \leftrightarrow \Box(p \wedge q)$ and $\Box\top$ and is closed under *congruence*: if $A \leftrightarrow B \in \mathcal{L}$, then $\Box A \leftrightarrow \Box B \in \mathcal{L}$.

Example 2.1.6.
1. The set of theorems of classical (resp. intuitionistic resp. minimal) logic itself is a $\{\to, \wedge, \vee, \bot\}$-logic based on classical (resp. intuitionistic resp. minimal) logic.
2. The standard modal logics such as $\mathsf{K}, \mathsf{KT}, \mathsf{K4}$ and $\mathsf{S4}$ (see e.g. [HC96, BdRV01]) are normal $\Lambda_\Box$-logics based on classical propositional logic.

For more details on classical, intuitionistic of minimal propositional logic see e.g. [TS00]. In order to make the roles of certain structural rules precise we will follow the standard procedure [NvP01, TS00] and treat sequents over a set F of formulae in terms of finite multisets.

Definition 2.1.7. Formally, for a set F a *finite multiset over* F is a function $\Gamma : F \to \mathbb{N}$ with finite support. If Γ is a multiset over F, then for elements A of F we say that A is an *element of* Γ and write $A \in \Gamma$, if $\Gamma(A) > 0$. Similarly, we extend the set theoretic notion of union to multisets: if Γ and Δ are multisets over F, then $\Gamma \cup \Delta$ is the multiset defined by $(\Gamma \cup \Delta)(A) := \Gamma(A) + \Delta(A)$ for all $A \in F$. We also often write Γ, Δ for the union $\Gamma \cup \Delta$. If Γ is a multiset, then the *support of* Γ is the multiset $\mathsf{Supp}(\Gamma)$ defined by $\mathsf{Supp}(\Gamma)(A) = 0$ if $\Gamma(A) = 0$ and $\mathsf{Supp}(\Gamma)(A) = 1$ otherwise. We sometimes write $A \in \Gamma \cap \Delta$ for $A \in \Gamma$ and $A \in \Delta$. For a multiset Γ the *size of* Γ is defined as $|\Gamma| := \sum_{A \in \Gamma} \Gamma(A)$. Since every finite set can be viewed as the support of a finite multiset we also identify finite sets with multisets where every element occurs with multiplicity one. For a finite multiset $\Gamma = A_1, \dots, A_n$ we also write $\bigwedge \Gamma$ for $A_1 \wedge \ldots \wedge A_n$ and $\bigvee \Gamma$ for $A_1 \vee \cdots \vee A_n$. If Γ is empty as usual we set $\bigwedge \Gamma := \top$ and $\bigvee \Gamma := \bot$.

Definition 2.1.8. Let F be a set of formulae. A *symmetric sequent over* F is a tuple $\Gamma \Rightarrow \Delta$ of finite multisets of formulae in F. The set of all symmetric sequents over F is denoted by $\mathcal{S}(F)$. An *asymmetric sequent over* F is a symmetric sequent $\Gamma \Rightarrow \delta$ over F, where δ is subject to the cardinality restriction $|\delta| \leq 1$. The *size* of a sequent $\Gamma \Rightarrow \Delta$ is defined as $|\Gamma \Rightarrow \Delta| := |\Gamma| + |\Delta|$. Occasionally we slightly abuse notation and given two sequents $\Gamma_1 \Rightarrow \Gamma_2$ and $\Sigma_1 \Rightarrow \Sigma_2$ write $\Gamma_1 \Rightarrow \Gamma_2 \subseteq \Sigma_1 \Rightarrow \Sigma_2$ if for $i = 1, 2$ and every formula A we have $\Gamma_i(A) \leq \Sigma_i(A)$.

$\dfrac{p_1, \dots, p_n \Rightarrow q}{\Box p_1, \dots, \Box p_n \Rightarrow \Box q}\ (\Rightarrow \Box)_1$	$\dfrac{\Gamma, p \Rightarrow \Delta}{\Gamma, \Box p \Rightarrow \Delta}\ (\Box \Rightarrow)_0$	$\dfrac{\Box\Gamma \Rightarrow p}{\Box\Gamma \Rightarrow \Box p}\ (\Rightarrow \Box)_0$

Table 2.1: Some standard sequent rules for modal logics taken from [Wan02]. The notation is slightly adapted.

Example 2.1.9. We consider sequents over the set $\mathcal{F}(\Lambda_\Box)$ of formulae.

1. For the two symmetric sequents $\Gamma \Rightarrow \Delta := \Box(p \wedge q), q \vee r \Rightarrow s, s, r \to \bot$ and $\Sigma \Rightarrow \Pi := \Box(p \wedge q) \Rightarrow$ we have $|\Gamma \Rightarrow \Delta| = 5$ and $|\Sigma \Rightarrow \Pi| = 1$. Furthermore since the formula $\Box(p \wedge q)$ occurs on the left hand side of $\Gamma \Rightarrow \Delta$ we have $\Sigma \Rightarrow \Pi \subseteq \Gamma \Rightarrow \Delta$.

2. For the two asymmetric sequents $\Gamma \Rightarrow \delta := \Box(p \wedge q), q \vee r, q \vee r \Rightarrow s$ and $\Sigma \Rightarrow \pi := \Box(p \wedge q), \Box(p \wedge q), q \vee r \Rightarrow s$ we have $|\Gamma \Rightarrow \delta| = |\Sigma \Rightarrow \pi| = 4$, but neither $\Gamma \Rightarrow \delta \subseteq \Sigma \Rightarrow \pi$ nor $\Sigma \Rightarrow \pi \subseteq \Gamma \Rightarrow \delta$.

Symmetric and asymmetric sequents are also known in the literature as multi- and single-succedent sequents respectively. In the following we will develop our theory for the symmetric and asymmetric frameworks in parallel. Thus unless otherwise stated all definitions and results apply to both frameworks, where in the asymmetric case we silently impose the cardinality restriction on every sequent. Later we will use the symmetric framework for modal logics based on classical propositional logic and the asymmetric framework for those based on minimal or intuitionistic propositional logic.

2.2 Intuitions

Before futher developing our general framework for sequent calculi and cut elimination we briefly pause and consider some concrete examples which showcase the intuitions behind the technical definitions. More precisely we are going to look at the modal rules for the well-known modal logics K, KT and $\mathsf{S4}$, which already introduce the concepts fundamental for the general framework. Since this section is concerned only with the intuitions we do not concern ourselves with precise definitions – these are given in the subsequent sections. The standard sequent rules governing the behaviour of the modality $\Box$ in these modal logics as found in the literature [Wan02] are given with slightly adapted notation in Table 2.1. Precise definitions of these rules and the notion of a sequent rule itself will be given in Section 2.3. To construct a sequent calculus for modal logic K we need to add the rule $(\Rightarrow \Box)_1$ to a standard sequent calculus for classical propositional logic. For now we do not consider the propositional rules, again they will be examined in the following section. For the modal logic KT we need to add both the rule $(\Rightarrow \Box)_1$ and the rule $(\Box \Rightarrow)_0$, and finally for the logic $\mathsf{S4}$ we add all three of the rules $(\Rightarrow \Box)_1$, $(\Box \Rightarrow)_0$ and $(\Rightarrow \Box)_0$. While in the last case it would also suffice to add only $(\Box \Rightarrow)_0$ and $(\Rightarrow \Box)_0$ for illustrative purposes we consider the calculus with all three rules.

If we take a closer look at these three rules we see the common characteristic that when passing from the premiss (at the top) to the conclusion (at the bottom) they introduce a single layer of modalities but otherwise keep all involved formulae intact and on the same side of the sequent. On the other hand they differ in the way they deal with the context, i.e., those formulae which occur unchanged both in the premiss and the conclusion. The rule $(\Rightarrow \Box)_1$ is very simple in this respect since it does not involve any context at all. This rule is an example of what we will later call *one-step rules*. The rule $(\Box \Rightarrow)_0$ on the other hand allows for a context, but without restricting it at all and is an example of a *shallow rule*. It should be noted that the structure of this rule is the same as that of the propositional rules, so the following remarks about this rule carry over to the propositional rules as well. Finally, the context in rule $(\Rightarrow \Box)_0$ is restricted to boxed formulae on the left hand side of the sequent. To handle this kind of rules we will introduce the notion of a *context restriction*. These three kinds of sequent rules will be defined precisely in Definitions 2.3.3 and 2.3.6 and will form the fundamental notions for the rest of our investigations.

Since logics in the sense of Definition 2.1.5 are closed under modus ponens, one way of constructing a complete sequent calculus for such a logic is to initially consider the sequent calculus with an additional rule called the *cut rule*. This rule has the form

$$\frac{\Gamma \Rightarrow A, \Delta \qquad \Sigma, A \Rightarrow \Pi}{\Gamma, \Sigma \Rightarrow \Delta, \Pi}\ \mathsf{Cut}$$

and, while allowing us to simulate modus ponens, is rather unpleasant in that it involves a formula, called the *cut formula*, which occurs in its premisses but not in its conclusion. This makes the uncontrolled use of the cut rule unsuitable for the purpose of a decision procedure for the logic under consideration since in derivations using this rule the occurring formulae may not be part of the conclusion. Thus the next and very important step is to show that the cut rule can be eliminated from the calculus, i.e., that every sequent which is derivable using the cut rule is also derivable without using this rule.

The main intuition behind this procedure of cut elimination is that applications of the cut rule are permuted upwards in a derivation until they arrive at the leaves, where they usually are easily replaced by a different rule. Wherever it is not possible to permute the application of the cut rule upwards it is replaced by an application of the cut rule where the cut formula is of lower complexity than the cut formula of the original cut. Then by induction on the complexity of the cut formula we eliminate all such cuts. Thus for example in the calculus for modal logic K the situation

$$\frac{\dfrac{\begin{matrix}\vdots\\ A_1, \dots, A_n \Rightarrow B\end{matrix}}{\Box A_1, \dots, \Box A_n \Rightarrow \Box B}\ (\Rightarrow \Box)_1 \qquad \dfrac{\begin{matrix}\vdots\\ B, C_1, \dots, C_m \Rightarrow D\end{matrix}}{\Box B, \Box C_1, \dots, \Box C_m \Rightarrow \Box D}\ (\Rightarrow \Box)_1}{\Box A_1, \dots, \Box A_n, \Box C_1, \dots, \Box C_m \Rightarrow \Box D}\ \mathsf{Cut}$$

is transformed into

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ A_1,\ldots,A_n \Rightarrow B\end{array} \qquad \begin{array}{c}\vdots\\ B, C_1,\ldots,C_m \Rightarrow D\end{array}}{A_1,\ldots,A_n, C_1,\ldots,C_m \Rightarrow D}\,\mathsf{Cut}}{\Box A_1,\ldots,\Box A_n, \Box C_1,\ldots,\Box C_m \Rightarrow \Box D}\,(\Rightarrow \Box)_1$$

Here the original application of the cut rule with cut formula $\Box B$ is replaced by an application of the cut rule with cut formula B. This new cut is both further up in the derivation and on a formula with lower complexity than the original cut formula. Note that in order to be able to perform this transformation we need to be able to apply the rule $(\Rightarrow \Box)_1$ to the sequent $A_1,\ldots,A_n, C_1,\ldots,C_m \Rightarrow D$. In general this is a property of the set of sequent rules and not always possible. We will investigate this further in Section 2.4.

For rules such as the rule $(\Box \Rightarrow)_0$ in the sequent calculus for modal logic KT or the propositional rules which involve a context we might encounter the new situation that the cut formula is part of the context. In this case the cut is pushed upwards into the premisses of the corresponding rule. Thus e.g. the situation:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ \Gamma, A \Rightarrow \Delta, \Box B\end{array}}{\Gamma, \Box A \Rightarrow \Delta, \Box B}\,(\Box \Rightarrow)_0 \qquad \dfrac{\begin{array}{c}\vdots\\ B, C_1,\ldots,C_n \Rightarrow D\end{array}}{\Box B, \Box C_1,\ldots,\Box C_n \Rightarrow \Box D}\,(\Rightarrow \Box)_1}{\Gamma, \Box A, \Box C_1,\ldots,\Box C_n \Rightarrow \Delta, \Box D}\,\mathsf{Cut}$$

where the cut formula $\Box B$ is part of the context of the rule $(\Box \Rightarrow)_0$ is transformed into

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ \Gamma, A \Rightarrow \Delta, \Box B\end{array} \qquad \dfrac{\begin{array}{c}\vdots\\ B, C_1,\ldots,C_n \Rightarrow D\end{array}}{\Box B, \Box C_1,\ldots,\Box C_n \Rightarrow \Box D}\,(\Rightarrow \Box)_1}{\Gamma, A, \Box C_1,\ldots,\Box C_n \Rightarrow \Delta, \Box D}\,\mathsf{Cut}}{\Gamma, \Box A, \Box C_1,\ldots,\Box C_n \Rightarrow \Delta, \Box D}\,(\Box \Rightarrow)_0$$

Here the new application of the cut rule is one step closer to the leaves of the derivation. The situation that the cut formula is part of the context in both of the last applied rules is similar. Note that since arbitrary formulae are allowed to occur in the context of rules like $(\Box \Rightarrow)_0$ we can always perform such a transformation if the cut formula is part of the context of at least one of the last applied rules. This changes if we consider rules with restricted context such as the rule $(\Rightarrow \Box)_0$ for modal logic $\mathsf{S4}$. In this case we might encounter the following situation:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ A_1,\ldots,A_n \Rightarrow B\end{array}}{\Box A_1,\ldots,\Box A_n \Rightarrow \Box B}\,(\Rightarrow \Box)_1 \qquad \dfrac{\begin{array}{c}\vdots\\ \Box B, \Box\Gamma \Rightarrow C\end{array}}{\Box B, \Box\Gamma \Rightarrow \Box C}\,(\Rightarrow \Box)_0}{\Box\Gamma, \Box A_1,\ldots,\Box A_n \Rightarrow \Box C}\,\mathsf{Cut}$$

This is transformed into

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ A_1,\dots,A_n \Rightarrow B\end{array}}{\Box A_1,\dots,A_n \Rightarrow \Box B}\,(\Rightarrow\Box)_1 \qquad \begin{array}{c}\vdots\\ \Box B, \Box\Gamma \Rightarrow C\end{array}}{\dfrac{\Box\Gamma, \Box A_1,\dots,\Box A_n \Rightarrow C}{\Box\Gamma, \Box A_1, \Box A_n \Rightarrow \Box C}\,(\Rightarrow\Box)_0}\,\mathsf{Cut}$$

where again the newly introduced cut is one step closer to the leaves of the derivation. In this case we can apply the rule $(\Rightarrow \Box)_0$ below the cut only because the formulae $\Box A_1, \dots, \Box A_n$ introduced by the rule $(\Rightarrow \Box)_1$ are boxed and therefore satisfy the restriction imposed on the context in rule $(\Rightarrow \Box)_0$. For arbitrary rule sets this can not always be done. Again a similar situation arises if the cut formula is part of the context in both last applied rules above the cut. We will see a formalisation of these intuitions in Section 2.4.

One important application of cut-free sequent calculi is their use in decision procedures for the corresponding logics. The main intuition here is that if all of the logical rules of the sequent calculus have the *subformula property*, i.e., if the premisses of applications of these rules only contain subformulae of their conclusions, then the number of different formulae possibly relevant for a derivation of a given sequent is bounded. Of course the cut rule does not satisfy this property, but often the rules of a cut-free calculus do. In particular we will only consider logical rules which satisfy the subformula property. Under some additional assumptions this means that also the number of different sequents possibly relevant for such a derivation is bounded. In many cases this is enough to obtain a decision procedure for the logic, either by enumerating all the derivable sequents possibly occurring in a derivation of a given sequent or by employing the method of *backwards proof search*. While in our case the first method in general necessitates the enumeration of exponentially many sequents and thus typically results in a decision procedure of exponential time complexity, backwards proof search often can be done depth first and thus due to the specific format of the rules often results in a procedure of polynomial space complexity. We will consider a detailed treatment of these issues in Section 2.7. But let us first turn to the precise definitions of sequent rules.

2.3 Rules with Context Restrictions

As mentioned in the previous section when looking at standard sequent systems for most modal logics such as $\mathsf{K}, \mathsf{KT}, \mathsf{KD}$ or $\mathsf{S4}$ as given e.g. in [Wan02] we notice two main features of the logical rules (again consider Table 2.1 on p. 28 for examples). First, we have a number of *principal formulae* in the conclusion, which are stripped of one layer of modalities when passing over to the premisses. The latter point is particularly interesting from the perspective of backwards proof search, since it can be used to ensure termination of the procedure. Second, we might have a *context*, that is a number of formulae which are not changed when passing from conclusion to premisses. These context formulae can be arbitrary, as in rule $(\Box \Rightarrow)_0$, or they can be *restricted* to a certain component of the sequent or a certain format, e.g. to boxed formulae in the left component as in rule $(\Rightarrow \Box)_0$. In order to make these distinctions precise, we introduce the notion of *context restrictions*:

Definition 2.3.1. Let F be a set of formulae. A *context restriction over* F is a tuple $\langle F_1, F_2 \rangle$ of sets of formulae in F. The set F_1 is called the *left component* of the restriction. Analogously, F_2 is called the *right component*. For a (symmetric or asymmetric) sequent $\Gamma \Rightarrow \Delta$ and a context restriction $\mathcal{C} = \langle F_1, F_2 \rangle$ the *restriction of* $\Gamma \Rightarrow \Delta$ *according to* $\mathcal{C}$ is the sequent $(\Gamma \Rightarrow \Delta) \upharpoonright_{\mathcal{C}} := \Gamma \upharpoonright_{F_1} \Rightarrow \Delta \upharpoonright_{F_2}$, where $\Gamma \upharpoonright_{F_1}$ is the multiset Γ restricted to substitution instances of formulae in F_1. We say that a sequent $\Gamma \Rightarrow \Delta$ *satisfies the restriction* $\mathcal{C}$ if $(\Gamma \Rightarrow \Delta) \upharpoonright_{\mathcal{C}} = \Gamma \Rightarrow \Delta$ and write $\mathfrak{C}(F)$ for the set of context restrictions over F. If $\mathcal{C}_1$ and $\mathcal{C}_2$ are context restrictions we say that the restriction $\mathcal{C}_1$ *satisfies the restriction* $\mathcal{C}_2$ if for every sequent $\Gamma \Rightarrow \Delta$ we have $(\Gamma \Rightarrow \Delta) \upharpoonright_{\mathcal{C}_1} \upharpoonright_{\mathcal{C}_2} = (\Gamma \Rightarrow \Delta) \upharpoonright_{\mathcal{C}_1}$.

Example 2.3.2. The following are some often encountered context restrictions which we will also use later on.

1. The *trivial restriction* $\mathcal{C}_{\mathsf{id}} := \langle \{p\}, \{p\} \rangle$ poses no restriction at all to the sequents since every formula is a substitution instance of the formula p. Thus for every sequent we have $(\Gamma \Rightarrow \Delta) \upharpoonright_{\mathcal{C}_{\mathsf{id}}} = \Gamma \Rightarrow \Delta$. Every context restriction satisfies the trivial restriction.

2. The *empty restriction* $\mathcal{C}_{\emptyset} := \langle \emptyset, \emptyset \rangle$ deletes every formula in a sequent since no formula is a substitution instance of a formula in $\emptyset$. For every sequent we have $(\Gamma \Rightarrow \Delta) \upharpoonright_{\mathcal{C}_{\emptyset}} = \;\Rightarrow\;$. The empty restriction satisfies every context restriction.

3. The restriction $\mathcal{C}_{\mathsf{4}} := \langle \{\Box p\}, \emptyset \rangle$ restricts the left hand side of a sequent to boxed formulae and deletes the every formula on the right hand side. E.g we have $(q, C \wedge D, \Box(A \vee B) \Rightarrow \Box D, p) \upharpoonright_{\mathcal{C}_{\mathsf{4}}} = \Box(A \vee B) \Rightarrow$.

4. The restriction $\mathcal{C}_{\mathsf{45}} := \langle \{\Box p\}, \{\Box p\} \rangle$ restricts both left and right hand side of a sequent to boxed formulae: we have $(q, C \wedge D, \Box(A \vee B) \Rightarrow \Box D, p) \upharpoonright_{\mathcal{C}_{\mathsf{45}}} = \Box(A \vee B) \Rightarrow \Box D$. The restriction $\mathcal{C}_{\mathsf{4}}$ satisfies the restriction $\mathcal{C}_{\mathsf{45}}$ but not vice versa since for the sequent $\Rightarrow \Box p$ we have $(\;\Rightarrow \Box p) \upharpoonright_{\mathcal{C}_{\mathsf{45}}} \upharpoonright_{\mathcal{C}_{\mathsf{4}}} = \;\Rightarrow\; \neq \;\Rightarrow \Box p = (\;\Rightarrow \Box p) \upharpoonright_{\mathcal{C}_{\mathsf{45}}}$.

Using the notion of a context restriction we generalise the above given examples of modal sequent rules to the format of rules with context restrictions.

Definition 2.3.3. A *rule with context restrictions* or simply a *rule* is a tuple $(\mathcal{P}/\Sigma \Rightarrow \Pi)$ where $\mathcal{P} \subseteq \mathcal{S}(\mathsf{Var}) \times \mathfrak{C}(\mathcal{F})$ is the set of *premisses* with associated context restrictions and $\Sigma \Rightarrow \Pi \in \mathcal{S}(\Lambda(\mathsf{Var}))$ are the *principal formulae.* We furthermore stipulate that every rule with context restrictions is subject to the *variable condition*, which states that no variable occurs more than once in the principal formulae and that every variable occurring in the premisses also occurs in the principal formulae. In the asymmetric case moreover we assume that for every premiss $(\Gamma \Rightarrow \delta; \mathcal{C})$ with $\delta \neq \emptyset$ the right component of $\mathcal{C}$ is empty. For a rule R we sometimes write $\mathsf{PF}(R)$ for the principal formulae of R. An *application* of a rule $R = (\mathcal{P}/\Sigma \Rightarrow \Pi)$ is given by a substitution $\sigma : \mathsf{Var} \to \mathcal{F}$ and a *context* $\Theta \Rightarrow \Xi \in \mathcal{S}(\mathcal{F})$ and written as

$$\frac{\{\ \Theta \restriction_{\mathcal{C}}, \Gamma\sigma \Rightarrow \Delta\sigma, \Xi \restriction_{\mathcal{C}} \mid (\Gamma \Rightarrow \Delta; \mathcal{C}) \in \mathcal{P}\ \}}{\Theta, \Sigma\sigma \Rightarrow \Pi\sigma, \Xi}\ R$$

We call the sequent $\Theta, \Sigma\sigma \Rightarrow \Pi\sigma, \Xi$ the *conclusion* and the sequents $\Theta \restriction_{\mathcal{C}}, \Gamma\sigma \Rightarrow \Delta\sigma, \Xi \restriction_{\mathcal{C}}$ the *premisses* of the application of R. Thus the variables in the principal formulae and the active part of the premisses are substituted by formulae, and each premiss carries over the context restricted according to its associated context restriction.

Example 2.3.4. Consider the set $\Lambda = \{\heartsuit\}$ of connectives. We take as an example the rule

$$R = \{(\ \Rightarrow q; \langle\{\heartsuit p\}, \emptyset\rangle), (p \Rightarrow\ ; \mathcal{C}_\emptyset)\}/\heartsuit p \Rightarrow \heartsuit q\ .$$

The context restriction $\langle\{\heartsuit p\}, \emptyset\rangle$ copies over all formulae of the form $\heartsuit A$ from the context. Thus an application of this rule is given e.g. by the substitution σ with $\sigma(p) = B$ and $\sigma(q) = C$ and the context $\Gamma, \heartsuit A_1, \ldots, \heartsuit A_n \Rightarrow \Delta$, where no formula in Γ has the form $\heartsuit A$ and is written as

$$\frac{\heartsuit A_1, \ldots, \heartsuit A_n \Rightarrow C \qquad B \Rightarrow}{\Gamma, \heartsuit A_1, \ldots, \heartsuit A_n, \heartsuit B \Rightarrow \heartsuit C, \Delta}$$

If the boolean operators are in the set of connectives it will be convenient to distinguish purely modal from mixed rules.

Definition 2.3.5. A rule with context restrictions $R = \mathcal{P}/\Sigma \Rightarrow \Pi$ is a *modal rule* if no boolean connectives occur in its principal formulae $\Sigma \Rightarrow \Pi$.

In the following we will also consider two slightly smaller classes of sequent rules. These are the result of limiting the context restrictions occurring in the premisses in different ways. The first format, that of *one-step* rules stems from [PS08] and does not allow the propagation of any context formulae, while the format of *shallow rules* is a natural extension which already captures the standard propositional rules.

Definition 2.3.6. A rule with restrictions is called a *one-step rule* if the only restriction occurring in it is $\mathcal{C}_\emptyset$ and a *shallow rule* if all its restrictions are either $\mathcal{C}_\emptyset$ or $\mathcal{C}_{\mathsf{id}}$. In the asymmetric case for shallow rules we also allow the restriction $\langle\{p\}, \emptyset\rangle$ for premisses $(\Gamma \Rightarrow \delta; \mathcal{C})$ with $\delta \neq \emptyset$.

Example 2.3.7.
1. The set Gc of rules of classical propositional logic as given in Table 2.2 as well as the sets Gi and Gm of rules for intuitionistic resp. minimal logic as given in Table 2.3 can be seen as sets of rules with restrictions. Since the only context restriction

$$\frac{}{\Gamma, \bot \Rightarrow \Delta}\ \bot_{\mathsf{L}}$$

$$\frac{\Gamma \Rightarrow A, \Delta \quad \Gamma \Rightarrow B, \Delta}{\Gamma \Rightarrow A \wedge B, \Delta}\ \wedge_{\mathsf{R}} \qquad \frac{\Gamma \Rightarrow A, B, \Delta}{\Gamma \Rightarrow A \vee B, \Delta}\ \vee_{\mathsf{R}} \qquad \frac{\Gamma, A \Rightarrow B, \Delta}{\Gamma \Rightarrow A \to B, \Delta}\ \to_{\mathsf{R}}$$

$$\frac{\Gamma, A, B \Rightarrow \Delta}{\Gamma, A \wedge B \Rightarrow \Delta}\ \wedge_{\mathsf{L}} \qquad \frac{\Gamma, A \Rightarrow \Delta \quad \Gamma, B \Rightarrow \Delta}{\Gamma, A \vee B \Rightarrow \Delta}\ \vee_{\mathsf{L}} \qquad \frac{\Gamma, \Rightarrow A, \Delta \quad \Gamma, B \Rightarrow \Delta}{\Gamma, A \to B \Rightarrow \Delta}\ \to_{\mathsf{L}}$$

Table 2.2: The rules in Gc.

$$\frac{}{\Gamma, \bot \Rightarrow \delta}\ \bot_{\mathsf{L}}$$

$$\frac{\Gamma \Rightarrow A \quad \Gamma \Rightarrow B}{\Gamma \Rightarrow A \wedge B}\ \wedge_{\mathsf{R}} \qquad \frac{\Gamma \Rightarrow A_i}{\Gamma \Rightarrow A_1 \vee A_2}\ \vee_{\mathsf{R}}\ (i = 1, 2) \qquad \frac{\Gamma, A \Rightarrow B}{\Gamma \Rightarrow A \to B}\ \to_{\mathsf{R}}$$

$$\frac{\Gamma, A, B \Rightarrow \delta}{\Gamma, A \wedge B \Rightarrow \delta}\ \wedge_{\mathsf{L}} \qquad \frac{\Gamma, A \Rightarrow \delta \quad \Gamma, B \Rightarrow \delta}{\Gamma, A \vee B \Rightarrow \delta}\ \vee_{\mathsf{L}} \qquad \frac{\Gamma, \Rightarrow A \quad \Gamma, B \Rightarrow \delta}{\Gamma, A \to B \Rightarrow \delta}\ \to_{\mathsf{L}}$$

Table 2.3: The rules in Gi. For Gm we drop the rule $\bot_{\mathsf{L}}$ from Gi.

occurring in the rules is $\mathcal{C}_{\mathsf{id}}$ they are in particular shallow rules. Table 2.4 gives the rules in our notation for a few examples.

2. For every n-ary modality $\heartsuit \in \Lambda$ the *congruence rule* for $\heartsuit$ given by $\mathsf{Cong}_{\heartsuit} = \{(p_i \Rightarrow q_i; \mathcal{C}_\emptyset) \mid 1 \leq i \leq n\} \cup \{(q_i \Rightarrow p_i; \mathcal{C}_\emptyset) \mid 1 \leq i \leq n\}/\heartsuit(p_1, \dots, p_n) \Rightarrow \heartsuit(q_1, \dots, q_n)$ and the *monotonicity rule* for $\heartsuit$ given by $\mathsf{Mon}_{\heartsuit} = \{(p_i \Rightarrow q_i; \mathcal{C}_\emptyset) \mid 1 \leq i \leq n\}/\heartsuit(p_1, \dots, p_n) \Rightarrow \heartsuit(q_1, \dots, q_n)$ are rules with restrictions. Since the restriction $\mathcal{C}_\emptyset$ is the only restriction occurring in these rules they are also one-step rules (and shallow rules).

3. The modal rules $\mathsf{K}_n, R_{\mathsf{T}}, \mathsf{4}_n$ and R_{45} given in Table 2.4 are rules with restrictions. Rules K_n and R_{T} are shallow rules, while K_n also is a one-step rule. (A brief note on notation: we use subscripts for variable numbers of principal formulae. For rules which are the result of translating an axiom (A) into a rule we write R_{A}.)

4. The modal rule
$$\frac{\Box\Gamma, \Box B \Rightarrow B}{\Box\Gamma \Rightarrow \Box B}$$
from [Lei81], which is used in sequent style presentations of Gödel-Löb logic, cannot be seen as a rule with context restrictions, since the formula $\Box B$ occurs in its premiss and thus cannot be principal in its conclusion, but on the other hand it changes the side of the sequent and thus is not a context formula.

5. The modal rule
$$\frac{A_1, \dots, A_n \Rightarrow B, \Box C_1, \dots, \Box C_m}{\Box A_1, \dots, \Box A_n \Rightarrow \Box B, C_1, \dots, C_m}$$

$\dfrac{}{\Gamma, \bot \Rightarrow \Delta}\ \bot_\mathsf{L}$	$\emptyset / \bot \Rightarrow$
$\dfrac{\Gamma \Rightarrow A, \Delta \quad \Gamma \Rightarrow B, \Delta}{\Gamma \Rightarrow A \wedge B, \Delta}\ \wedge_\mathsf{R}$	$\{(\ \Rightarrow p; \mathcal{C}_\mathsf{id}), (\ \Rightarrow q; \mathcal{C}_\mathsf{id})\} /\ \Rightarrow p \wedge q$
$\dfrac{A_1, \dots, A_n \Rightarrow B}{\Gamma, \Box A_1, \dots, \Box A_n \Rightarrow \Box B, \Delta}\ \mathsf{K}_n$	$\{(p_1, \dots, p_n \Rightarrow q;\ \mathcal{C}_\emptyset)\} / \Box p_1, \dots, \Box p_n \Rightarrow \Box q$
$\dfrac{\Gamma, A \Rightarrow \Delta}{\Gamma, \Box A \Rightarrow \Delta}\ R_\mathsf{T}$	$\{(p \Rightarrow\ ;\ \mathcal{C}_\mathsf{id})\} / \Box p \Rightarrow$
$\dfrac{\Box\Gamma, A_1, \dots, A_n \Rightarrow B}{\Sigma, \Box\Gamma, \Box A_1, \dots, \Box A_n \Rightarrow \Box B, \Delta}\ \mathsf{4}_n$	$\{(p_1, \dots, p_n \Rightarrow q;\ \mathcal{C}_\mathsf{4})\} / \Box p_1, \dots, \Box p_n \Rightarrow \Box q$
$\dfrac{\Box\Gamma \Rightarrow A, \Box\Delta}{\Sigma, \Box\Gamma \Rightarrow \Box A, \Box\Delta, \Pi}\ R_\mathsf{45}$	$\{(\ \Rightarrow p; \mathcal{C}_\mathsf{45})\} /\ \Rightarrow \Box p$

Table 2.4: Some standard modal rules in the notation as rules with context restrictions

$\dfrac{\Gamma \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta}\ \mathsf{W}_\mathsf{L}$	$\dfrac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow A, \Delta}\ \mathsf{W}_\mathsf{R}$
$\dfrac{\Gamma, A, A \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta}\ \mathsf{Con}_\mathsf{L}$	$\dfrac{\Gamma \Rightarrow B, B, \Delta}{\Gamma \Rightarrow B, \Delta}\ \mathsf{Con}_\mathsf{R}$
$\dfrac{\Gamma \Rightarrow A, \Delta \quad \Sigma, A \Rightarrow \Pi}{\Gamma, \Sigma \Rightarrow \Delta, \Pi}\ \mathsf{Cut}$	$\dfrac{\Gamma, \Gamma, \Sigma \Rightarrow \Delta, \Delta, \Pi}{\Gamma, \Sigma \Rightarrow \Delta, \Pi}\ \mathsf{Mcon}$

Table 2.5: The structural rules.

which is often used to capture modal logic B (see e.g. [Tak92, Wan02]) also cannot be seen as a rule with context restrictions, since this format does not allow the modal rank to decrease when passing from the premisses to the conclusion of a rule.

Notation 2.3.8. If a result holds for the rule set Gc in the symmetric framework as well as for the rule sets Gi and Gm in the asymmetric framework, then we also write $\mathsf{G[cim]}$ and similarly for $\mathsf{G[ci]}$. Furthermore, if a lemma or theorem presupposes the propositional rules we mark it with the corresponding letter $\mathsf{c}, \mathsf{i}, \mathsf{m}$ or a combination of these if it holds for the corresponding rule sets.

We also consider the standard structural rules.

Definition 2.3.9. The structural rules of *right-* (resp. *left-*) *Weakening* and *right-*(resp. *left-*) *Contraction* as well as *Cut* and *multi-Contraction* Mcon are given in Table 2.5. We also write W for $\{\mathsf{W}_\mathsf{L}, W_\mathsf{R}\}$ and Con for $\{\mathsf{Con}_\mathsf{L}, \mathsf{Con}_\mathsf{R}\}$. The notion of application of a rule is extended to the structural rules in the obvious way. For the contraction rules we also call the formula on which the contraction is performed the *contraction formula* and its occurrence in the conclusion the *principal formula* of the rule.

The structural rules are not rules with context restrictions in our sense since they do not introduce a layer of connectives in the principal formulae. The notions of a derivation and a derivable sequent are now defined in the standard way.

Definition 2.3.10. Let $\mathcal{R}$ be a set of structural rules and / or rules with context restrictions including the *axiom rules*

$$\frac{}{\Gamma, A \Rightarrow A, \Delta}\ \mathcal{A}$$

with *principal formulae* $A \Rightarrow A$ where A is an arbitrary formula and *context formulae* $\Gamma \Rightarrow \Delta$. A *derivation in* $\mathcal{R}$ is a finite tree where each node is labelled with a sequent such that for every node its sequent is the conclusion of an application of a rule in $\mathcal{R}$ and the sequents of the predecessors of this node are the premisses of this application. The *depth* of a derivation is the height of the underlying tree, that is the maximal length of a path from the root to a leaf. A sequent $\Gamma \Rightarrow \Delta$ is *derivable in* $\mathcal{R}$ if there is a derivation in $\mathcal{R}$ whose root is labelled with $\Gamma \Rightarrow \Delta$. If $\Gamma \Rightarrow \Delta$ is derivable in $\mathcal{R}$ we also write $\vdash_{\mathcal{R}} \Gamma \Rightarrow \Delta$. If $\mathcal{R}$ is a rule set and R a rule, then we often slightly abuse notation and write $\vdash_{\mathcal{R}R}$ instead of $\vdash_{\mathcal{R}\cup\{R\}}$. Furthermore, we write e.g. $\vdash_{\mathcal{R}[\mathsf{ConCutW}]}$ if the statement holds for the rule set $\mathcal{R}$ extended with an arbitrary subset of $\mathsf{ConCutW}$ and similarly for other sets of structural rules. If S is a set of sequents we say that a sequent $\Gamma \Rightarrow \Delta$ is *derivable from* S *in* $\mathcal{R}[\mathsf{ConCutW}]$ and write $S \vdash_{\mathcal{R}[\mathsf{ConCutW}]} \Gamma \Rightarrow \Delta$ if there is a derivation of $\Gamma \Rightarrow \Delta$ where the leafs are labelled with conclusions of applications of the axiom rule or sequents in S. A rule is *derivable* in a rule set $\mathcal{R}[\mathsf{ConCutW}]$ if for all its applications the conclusion is derivable from the set of its premisses in $\mathcal{R}[\mathsf{ConCutW}]$ and *admissible in* $\mathcal{R}[\mathsf{ConCutW}]$ if every sequent is derivable in $\mathcal{R}R[\mathsf{ConConW}]$ if and only if it is derivable in $\mathcal{R}[\mathsf{ConCutW}]$.

Convention: From now on unless stated otherwise whenever we talk about a set $\mathcal{R}$ of rules with restrictions we assume that $\mathcal{R}$ is closed under injective renamings of the variables and contains the axiom rules $\mathcal{A}$, the identity rule R_{id} and the congruence rules $\mathsf{Cong}_\heartsuit$ for every modality $\heartsuit \in \Lambda$.

As usual when presenting derivations we will sometimes abbreviate multiple applications of the same rule using double lines instead of single lines. Thus e.g. the derivation below left is abbreviated as shown below right.

$$\dfrac{\dfrac{\vdots \atop \Gamma, A, B, C, D \Rightarrow \Delta}{\Gamma, A, B, C \wedge D \Rightarrow \Delta}\ \wedge_{\mathsf{L}}}{\Gamma, A \wedge B, C \wedge D \Rightarrow \Delta}\ \wedge_{\mathsf{L}} \qquad\qquad \dfrac{\vdots \atop \Gamma, A, B, C, D \Rightarrow \Delta}{\Gamma, A \wedge B, C \wedge D \Rightarrow \Delta}\ \wedge_{\mathsf{L}}$$

The connection between logics and sequent calculi is given by the standard notions of soundness and completeness.

Definition 2.3.11. Let Λ be a set of connectives including the boolean connectives. Furthermore, let $\mathcal{R}$ be a set of rules with context restrictions and let $\mathcal{L}$ be a logic based on classical, intuitionistic or minimal propositional logic. We say that the sequent calculus given by $\mathcal{R}[\mathsf{CutConW}]$ is *sound for* $\mathcal{L}$ if for all sequents $\Gamma \Rightarrow \Delta$ we have

$$\vdash_{\mathcal{R}[\mathsf{CutConW}]} \Gamma \Rightarrow \Delta \quad \text{implies} \quad \bigwedge \Gamma \to \bigvee \Delta \in \mathcal{L} \ .$$

Similarly, we say that the sequent calculus given by $\mathcal{R}[\mathsf{CutConW}]$ is *complete for* $\mathcal{L}$ if for every sequent $\Gamma \Rightarrow \Delta$ we have

$$\bigwedge \Gamma \to \bigvee \Delta \in \mathcal{L} \quad \text{implies} \quad \vdash_{\mathcal{R}[\mathsf{CutConW}]} \Gamma \Rightarrow \Delta \ .$$

Note that in the asymmetric framework the right hand side of the sequent and thus also the succedent of the implication contains at most one formula. If $\mathcal{R}[\mathsf{ConW}]$ is complete for $\mathcal{L}$, then we also explicitly mention the absence of the cut rule and say that the calculus is *cut free complete for* $\mathcal{L}$.

By virtue of the format of rules with restrictions we now immediately obtain our first result.

Lemma 2.3.12 (Admissibility of Weakening)**.** *Let* $\mathcal{R}$ *be a set of rules with restrictions and let* $\Gamma \Rightarrow \Delta$ *be a sequent. Then we have*

$$\vdash_{\mathcal{R}\mathsf{W}[\mathsf{CutCon}]} \Gamma \Rightarrow \Delta \quad \textit{iff} \quad \vdash_{\mathcal{R}[\mathsf{CutCon}]} \Gamma \Rightarrow \Delta$$

and the depth of the derivations is preserved.

Proof. Standard by induction on the depth of the derivation, using the fact that in applications of rules with restrictions the weakening can be pushed into all those premisses whose restrictions are satisfied by the weakening formula. □

Thus we may simply drop the rule of Weakening from our rule set without changing the set of derivable sequents. Nevertheless, for convenience we will sometimes make use of the weakening rule.

2.4 Cut Elimination

Arguably one of the most important properties of a sequent system is admissibility of the cut rule. While this can be shown by semantically driven arguments such as proving completeness of the system without the cut rule, often we are interested in a constructive method for transforming a derivation with the cut rule into one without it, the standard reference for this of course being Gentzen's original proof of Cut Elimination for first-order logic in [Gen34]. The idea here is to permute applications of the cut rule upwards in the derivation until only cuts involving conclusions of the axiom rule or the rule $\bot_{\mathsf{L}}$ remain. These cuts can then easily be eliminated. For the permutation steps, depending on whether the cut formula was principal in the last applied rules of both premisses of the application of cut or not, the cut is either replaced by cuts on proper subformulae of the original cut formula or permuted upwards into the premisses of a rule where the cut formula was contextual.

We are now going to see some criteria on sets of rules which ensure that a modal analogue of this proof goes through. The main idea is that applications of cut where the cut formula is principal in the last applied rules of both premisses of the cut can be absorbed into the rule set. This way such an application of cut can be replaced by a number of cuts on formulae of lower complexity and an application of a rule from the rule set. In order to state this formally we are going to construct this latter rule in a uniform way. To get an idea of how this works assume we have the rules

$$R_1 = \mathcal{P}_1/\Sigma_1 \Rightarrow \Pi_1, \heartsuit p \quad \text{and} \quad R_2 = \mathcal{P}_2/\heartsuit p, \Sigma_2 \Rightarrow \Pi_2 \ .$$

Applying cut to the conclusions of these two rules yields the sequent $\Sigma_1, \Sigma_2 \Rightarrow \Pi_1, \Pi_2$, which we will take to be the conclusion of the new rule. The easiest option for the premisses of the new rule would be to simply take $\mathcal{P}_1 \cup \mathcal{P}_2$, but unfortunately in this case we might end up with the variable p occurring in the premisses but not in the conclusion, a situation which is not allowed by the rule format. Fortunately this can be easily changed by intuitively performing all the possible cuts between the premisses on the variable p. This process is captured in the following definition.

Definition 2.4.1. If $\mathcal{P} \subseteq \mathcal{S}(\mathsf{Var}) \times \mathfrak{C}(\mathcal{F})$ is a set of premisses with context restrictions, then for $p \in \mathsf{Var}$ the *p-elimination of* $\mathcal{P}$ is the set

$$\begin{aligned} \mathcal{P} \ominus p := \ & \{ \ (\mathsf{Supp}\,(\Gamma, \Sigma) \Rightarrow \mathsf{Supp}\,(\Delta, \Pi)\,;\ \mathcal{C}_1 \cup \mathcal{C}_2) \mid (\Gamma \Rightarrow \Delta, p;\ \mathcal{C}_1) \in \mathcal{P},\ (p, \Sigma \Rightarrow \Pi;\ \mathcal{C}_2) \in \mathcal{P} \ \} \\ & \cup \{ \ (\Gamma \Rightarrow \Delta;\ \mathcal{C}) \mid (\Gamma \Rightarrow \Delta;\ \mathcal{C}) \in \mathcal{P},\ p \notin \Gamma, \Delta \ \} \ , \end{aligned}$$

where for restrictions $\mathcal{C}_1 = \langle F_1, F_2 \rangle$ and $\mathcal{C}_2 = \langle G_1, G_2 \rangle$ we write $\mathcal{C}_1 \cup \mathcal{C}_2$ for $\langle F_1 \cup G_1, F_2 \cup G_2 \rangle$. Iterated elimination of variables $\vec{p} = p_1, \ldots, p_n$ is denoted by $\mathcal{P} \ominus \vec{p}$. For rules $R = (\mathcal{P}_R; \Gamma \Rightarrow \Delta, \heartsuit\vec{p})$ and $Q = (\mathcal{P}_Q; \heartsuit\vec{p}, \Sigma \Rightarrow \Pi)$ the *cut between R and Q on $\heartsuit\vec{p}$* is the rule

$$\mathsf{cut}(R, Q, \heartsuit\vec{p}) := ((\mathcal{P}_R \cup \mathcal{P}_Q) \ominus \vec{p};\ \Gamma, \Sigma \Rightarrow \Delta, \Pi) \ .$$

If for two context restrictions $\mathcal{C}_1 = \langle F_1, F_2 \rangle$ and $\mathcal{C}_2 = \langle G_1, G_2 \rangle$ for $i = 1, 2$ every formula which is a substitution instance of a formula in G_i is also a substitution instance of a formula in F_i, we also abbreviate the restriction $\mathcal{C}_1 \cup \mathcal{C}_2$ to $\mathcal{C}_1$.

Example 2.4.2. Let $\mathcal{P} := \{(p, q \Rightarrow r;\ \mathcal{C}_\emptyset),\ (\Rightarrow p;\ \mathcal{C}_4),\ (p \Rightarrow q;\ \mathcal{C}_{\mathsf{id}})\}$. Then the variable eliminations for the variables p, q, r are as follows.

1. The p-elimination of $\mathcal{P}$ is $\mathcal{P} \ominus p = \{(q \Rightarrow r;\ \mathcal{C}_4),\ (\Rightarrow q;\ \mathcal{C}_{\mathsf{id}})\}$.
2. The q-elimination of $\mathcal{P}$ is $\mathcal{P} \ominus q = \{(p \Rightarrow r;\ \mathcal{C}_{\mathsf{id}}),\ (\Rightarrow p;\ \mathcal{C}_4)\}$.
3. The r-elimination of $\mathcal{P}$ is $\mathcal{P} \ominus r = \{(\Rightarrow p;\ \mathcal{C}_4),\ (\Rightarrow q;\ \mathcal{C}_{\mathsf{id}})\}$.

Thus for a set $\mathcal{P}$ of premisses and a variable p the p-elimination of $\mathcal{P}$ is obtained by "performing all possible cuts on p". The concept of a cut between rules is illustrated by the following example.

Example 2.4.3. 1. For $\mathcal{P} = \{(\Rightarrow p; \mathcal{C}_{\mathsf{id}}),\ (\Rightarrow q;\ \mathcal{C}_{\mathsf{id}}),\ (p, q \Rightarrow\ ;\ \mathcal{C}_{\mathsf{id}})\}$ we have $\mathcal{P} \ominus (p, q) = \{ \Rightarrow\ ;\ \mathcal{C}_{\mathsf{id}})\}$ and thus $\mathsf{cut}(\wedge_{\mathsf{R}}, \wedge_{\mathsf{L}}, p \wedge q)$ is the *identity rule* $R_{\mathsf{id}} = \{ \Rightarrow\ ;\ \mathcal{C}_{\mathsf{id}})\}/ \Rightarrow$.

2. For $\mathcal{P} = \{(p_1, p_2 \Rightarrow q_1;\ \mathcal{C}_\emptyset),\ (q_1, q_2 \Rightarrow r;\ , \mathcal{C}_\emptyset)\}$ we have $\mathcal{P} \ominus q = \{(p_1, p_2, q_2 \Rightarrow r;\ \mathcal{C}_\emptyset)\}$ and thus $\mathsf{cut}(\mathsf{K}_2, \mathsf{K}_2, \Box q_1)$ is the rule $\mathsf{K}_3 = \{(p_1, p_2, q_2 \Rightarrow r;\ \mathcal{C}_\emptyset)\}/\Box p_1, \Box p_2, \Box q_2 \Rightarrow \Box r$.

3. For $\mathcal{P} = \{(p_1, p_2 \Rightarrow q;\ \mathcal{C}_\emptyset),\ (\Rightarrow p_1;\ \mathcal{C}_4)\}$ we have $\mathcal{P} \ominus p_1 = \{(p_2 \Rightarrow q;\ \mathcal{C}_4)\}$ and thus $\mathsf{cut}(4_0, \mathsf{K}_2, \Box p_1)$ is the rule $4_1 = \{(p_2 \Rightarrow q;\ \mathcal{C}_4)\}/\Box p_2 \Rightarrow \Box q$.

As a convenient fact about variable elimination and cuts between rules we note that modulo Weakening and Contraction the order in which we eliminate the variables in the premisses is not important.

Lemma 2.4.4. *For a set $\mathcal{P}$ of premisses and a sequent $\Theta \Rightarrow \Xi$ write $\mathcal{P}(\Theta \Rightarrow \Xi)$ for the set*

$$\{\Theta \restriction_{F_1}, \Gamma \Rightarrow \Delta, \Xi \restriction_{F_2} \mid (\Gamma \Rightarrow \Delta; \langle F_1, F_2 \rangle) \in \mathcal{P}\}\ .$$

Let $\mathcal{P}$ be a set of premisses, let p, q be variables and let $\Theta \Rightarrow \Xi$ be a sequent. Then every sequent in $(\mathcal{P} \ominus p, q)(\Theta \Rightarrow \Xi)$ is derivable from the sequents in $(\mathcal{P} \ominus q, p)(\Theta \Rightarrow \Xi)$ using only Weakening and Contraction.

Proof. Let $\Sigma \Rightarrow \Pi$ be a sequent in $(\mathcal{P} \ominus p, q)(\Theta \Rightarrow \Xi)$. We show that there is a sequent in $(\mathcal{P} \ominus q, p)(\Theta \Rightarrow \Xi)$ such that $\Sigma \Rightarrow \Pi$ is derivable from this sequent using Contraction and Weakening. We consider all possible cases for the construction of the sequent $\Sigma \Rightarrow \Pi$.

1. $\Sigma \Rightarrow \Pi \in (\mathcal{P} \ominus p)(\Theta \Rightarrow \Xi)$. Then q does not occur in $\Sigma \Rightarrow \Pi$.
   a) $\Sigma \Rightarrow \Pi \in \mathcal{P}(\Theta \Rightarrow \Xi)$. Then p does not occur in $\Sigma \Rightarrow \Pi$ either, and thus $\Sigma \Rightarrow \Pi \in (\mathcal{P} \ominus q, p)(\Theta \Rightarrow \Xi)$.
   b) $\Sigma \Rightarrow \Pi = \Sigma_1, \Sigma_2 \Rightarrow \Pi_1, \Pi_2$ where $\Sigma_1 \Rightarrow \Pi_1, p \in (\mathcal{P} \ominus q)(\Theta \Rightarrow \Xi)$ and $p, \Sigma_2 \Rightarrow \Pi_2 \in (\mathcal{P} \ominus q)(\Theta \Rightarrow \Xi)$. Since q does not occur in $\Sigma_i \Rightarrow \Pi_i$ for $i = 1, 2$ we have $\Sigma_i \Rightarrow \Pi_i \in (\mathcal{P} \ominus q)(\Theta \Rightarrow \Xi)$ and thus $\Sigma \Rightarrow \Pi \in (\mathcal{P} \ominus q, p)(\Theta \Rightarrow \Xi)$.
2. $\Sigma \Rightarrow \Pi = \Sigma_1, \Sigma_2 \Rightarrow \Pi_1, \Pi_2$ where $\Sigma_1 \Rightarrow \Pi_1, q \in (\mathcal{P} \ominus p)(\Theta \Rightarrow \Xi)$ and $q, \Sigma_2 \Rightarrow \Pi_2 \in (\mathcal{P} \ominus p)(\Theta \Rightarrow \Xi)$.

a) $\Sigma_1 \Rightarrow \Pi_1, q \in \mathcal{P}(\Theta \Rightarrow \Xi)$ and $q, \Sigma_2 \Rightarrow \Pi_2 \in \mathcal{P}(\Theta \Rightarrow \Xi)$. Then p does not occur in these sequents and we have $\Sigma \Rightarrow \Pi \in (\mathcal{P} \ominus q, p)(\Theta \Rightarrow \Xi)$.

b) $\Sigma_1 \Rightarrow \Pi_1, q \in \mathcal{P}(\Theta \Rightarrow \Xi)$ and $q, \Sigma_2, \Rightarrow \Pi_2 = \Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2$ for $\Gamma_1 \Rightarrow \Delta_1, p \in \mathcal{P}(\Theta \Rightarrow \Xi)$ and $p, \Gamma_2 \Rightarrow \Delta_2 \in \mathcal{P}(\Theta \Rightarrow \Xi)$. Then first cutting on q, then cutting on p yields a sequent in $(\mathcal{P} \ominus q, p)(\Theta \Rightarrow \Xi)$. Now applying Contraction gives the original sequent.

c) $q, \Sigma_2 \Rightarrow \Pi_2 \in \mathcal{P}(\Theta \Rightarrow \Xi)$ and $\Sigma_1 \Rightarrow \Pi_1, q = \Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2$ for $\Gamma_1 \Rightarrow \Delta_1, p \in \mathcal{P}(\Theta \Rightarrow \Xi)$ and $p, \Gamma_2 \Rightarrow \Delta_2 \in \mathcal{P}(\Theta \Rightarrow \Xi)$. Similar to the previous case.

d) $\Sigma_1 \Rightarrow \Pi_1, q = \Gamma_{1,1}, \Gamma_{1,2} \Rightarrow \Delta_{1,1}, \Delta_{1,2}$ and $q, \Sigma_2 \Rightarrow \Pi_2 = \Gamma_{2,1}, \Gamma_{2,2} \Rightarrow \Delta_{2,1}, \Delta_{2,2}$ where all of the sequents

$$\Gamma_{1,1} \Rightarrow \Delta_{1,1}, p \qquad p, \Gamma_{1,2} \Rightarrow \Delta_{1,2} \qquad \Gamma_{2,1} \Rightarrow \Delta_{2,1}, p \qquad p, \Gamma_{2,2} \Rightarrow \Delta_{2,2}$$

are in $\mathcal{P}(\Theta \Rightarrow \Xi)$. Then we have $q \in \Delta_{1,1} \cup \Delta_{1,2}$ and $q \in \Gamma_{2,1} \cup \Gamma_{2,2}$.

i. $q \notin \Delta_{1,1}$ and $q \notin \Gamma_{2,2}$. Then $\Gamma_{1,1}, \Gamma_{2,1} \Rightarrow \Delta_{1,1}, \Delta_{2,2} \in (\mathcal{P} \ominus q, p)(\Theta \Rightarrow \Xi)$ and the original sequent follows using Weakening.

ii. $q \notin \Delta_{1,2}$ and $q \notin \Gamma_{2,1}$. As in the last case.

iii. $q \notin \Delta_{1,2}$ and $q \notin \Gamma_{2,2}$. Then $\Gamma_{1,1}, \Gamma_{2,1} \Rightarrow \Delta_{1,1}, \Delta_{2,1}, p \in (\mathcal{P} \ominus q)(\Theta \Rightarrow \Xi)$ and thus $\Gamma_{1,1}, \Gamma_{2,1}, \Gamma_{2,2} \Rightarrow \Delta_{1,2}, \Delta_{2,1}, \Delta_{2,2} \in (\mathcal{P} \ominus q, p)(\Theta \Rightarrow \Xi)$. The original sequent follows using Weakening.

iv. $q \notin \Delta_{1,1}$ and $q \notin \Gamma_{2,1}$. As in the last case.

v. $q \notin \Omega$ for Ω exactly one of the multisets $\Delta_{1,1},\ \Delta_{1,2},\ \Gamma_{2,1},\ \Gamma_{2,2}$. We consider the case $q \notin \Delta_{1,1}$ and $q \in \Delta_{1,2} \cap \Gamma_{2,1} \cap \Gamma_{2,2}$. Then $\Gamma_{1,1} \Rightarrow \Delta_{1,1}, p \in (\mathcal{P} \ominus q)(\Theta \Rightarrow \Xi)$ and $p, \Gamma_{1,2}, \Gamma_{2,2} \Rightarrow \Delta_{1,2}, \Delta_{2,2} \in (\mathcal{P} \ominus q)(\Theta \Rightarrow \Xi)$. Thus $\Gamma_{1,1}, \Gamma_{1,2}, \Gamma_{2,2} \Rightarrow \Delta_{1,1}, \Delta_{1,2}, \Delta_{2,2} \in (\mathcal{P} \ominus q, p)(\Theta \Rightarrow \Xi)$ and the original sequent follows using Weakening. The other cases are similar.

vi. $q \in \Delta_{1,1} \cap \Delta_{1,2} \cap \Gamma_{2,1} \cap \Gamma_{2,2}$. Then $\Gamma_{1,1}, \Gamma_{2,1} \Rightarrow \Delta_{1,1}, \Delta_{2,1}, p \in (\mathcal{P} \ominus q)(\Theta \Rightarrow \Xi)$ and $p, \Gamma_{1,2}, \Gamma_{2,2} \Rightarrow \Delta_{1,2}, \Delta_{2,2} \in (\mathcal{P} \ominus q)(\Theta \Rightarrow \Xi)$ and thus the original sequent is in $(\mathcal{P} \ominus q, p)(\Theta \Rightarrow \Xi)$.

□

Cuts between rules will also play a major role later when we consider the construction of rule sets satisfying the criteria for Cut Elimination. For this reason we notice another very convenient property of cuts between modal rules in a rule set which also includes the standard propositional rules.

Lemma 2.4.5 (Soundness of Cuts)**.** *Let $\mathcal{R}$ be a set of rules with restrictions such that $\mathsf{G[cim]} \subseteq \mathcal{R}$, and let $R_1, R_2 \in \mathcal{R}$. Then every cut between R_1 and R_2 is a derivable rule in $\mathcal{R}\mathsf{CutConW}$.*

Proof. For the asymmetric case consider a unary modality $\heartsuit$ and the cut $\mathsf{cut}(R_1, R_2, \heartsuit p)$ between two rules $R_1 = (\mathcal{P}_1; \Gamma \Rightarrow \heartsuit p)$ and $R_2 = (\mathcal{P}_2; \heartsuit p, \Delta \Rightarrow \alpha)$. By definition we have $\mathsf{cut}(R_1, R_2, \heartsuit p) = ((\mathcal{P}_1 \cup \mathcal{P}_2) \ominus p; \Gamma, \Delta \Rightarrow \alpha)$. Our goal is to replace an arbitrary application of $\mathsf{cut}(R_1, R_2, \heartsuit p)$ in a derivation by applications of R_1 and R_2 and an application of the cut rule. Suppose the combined premisses of the two rules are

$$\mathcal{P}_1 \cup \mathcal{P}_2 = \{(\Theta_i \Rightarrow p; \mathcal{C}_i^r) \mid i \in I\} \cup \{(p, \Upsilon_j \Rightarrow \beta_j; \mathcal{C}_j^\ell) \mid j \in J\} \cup \{(\Xi_k \Rightarrow \gamma_k; \mathcal{C}_k^n) \mid k \in K\} ,$$

where $p \notin \Xi_k, \gamma_k$ for all $k \in K$. Now consider an application

$$\frac{\{\Sigma \upharpoonright_{\mathcal{C}_i^\ell \cup \mathcal{C}_j^r}, (\Theta_i, \Upsilon_j)\sigma \Rightarrow \beta_j\sigma \mid i \in I, j \in J\} \ \cup \ \{\Sigma \upharpoonright_{\mathcal{C}_k^n}, \Xi_k\sigma \Rightarrow \gamma_k\sigma \mid k \in K\}}{\Sigma, (\Gamma, \Delta)\sigma \Rightarrow \alpha\sigma}$$

of the rule $\mathsf{cut}(R_1, R_2, p)$ in a derivation. We construct a substitution instance of p by setting

$$P := \bigvee_{i \in I} \left(\bigwedge \Sigma \upharpoonright_{\mathcal{C}_i^r} \wedge \bigwedge \Theta_i\sigma \right) .$$

Note that in case $I = \emptyset$ we have $P = \bot$. If the base logic is minimal propositional logic in this case we set $P := \bigwedge_{j \in J} \beta_j$. Then for every $i \in I$ we can derive $\Sigma \upharpoonright_{\mathcal{C}_i^r}, \Theta_i\sigma \Rightarrow P$ from axioms using the right conjunction and disjunction rules. Furthermore for every $i \in I$ and $j \in J$ we get $\Sigma \upharpoonright_{\mathcal{C}_j^\ell}, \Sigma \upharpoonright_{\mathcal{C}_i^r}, \Theta_i\sigma, \Upsilon_j\sigma \Rightarrow \beta_j\sigma$ from the premisses of the application using Weakening. Thus for every $j \in J$ we have $\Sigma \upharpoonright_{\mathcal{C}_j^\ell}, P, \Upsilon_j\sigma \Rightarrow \beta_j\sigma$ by left conjunction and disjunction. If $I = \emptyset$ in the intuitionistic case we use the rule $\bot_\mathsf{L}$ and in the minimal case the rules $\mathcal{A}, \wedge_\mathsf{L}$. Now we can apply the rules R_1 and R_2 to these premisses, apply cut to the conclusions with cut formula $\heartsuit P$ and contract duplicate context formulae to arrive at the conclusion of the application of the cut. If the modality has arity greater than 1 we iterate the process.

The symmetric case is treated similarly. Remember that for a formula A we write $\neg A$ as an abbreviation for the formula $A \to \bot$. Now again we take I to be the set of indices for which $(\Theta_i \Rightarrow p, \Omega_i; \mathcal{C}_i^r)$ occurs in the combined premisses of the two rules and set

$$P := \bigvee_{i \in I} \left(\bigwedge \Sigma \upharpoonright_{\mathcal{C}_i^r} \wedge \neg \bigvee \Pi \upharpoonright_{\mathcal{C}_i^r} \wedge \bigwedge \Theta_i\sigma \wedge \neg \bigvee \Omega_i\sigma \right)$$

where Π is the right hand part of the context of the rule application. Then again we can derive $\Sigma \upharpoonright_{\mathcal{C}_i^r}, \Theta_i\sigma \Rightarrow P, \Omega_i\sigma, \Pi \upharpoonright_{\mathcal{C}_i^r}$ from axioms and for every $j \in J$ the sequent $\Sigma \upharpoonright_{\mathcal{C}_j^\ell}, \Upsilon_j\sigma, P \Rightarrow \Xi\sigma, \Pi \upharpoonright_{\mathcal{C}_j^\ell}$ from the premisses of $\mathsf{cut}(R_1, R_2, p)$ using propositional logic. Now applications of R_1 and R_2, a cut and contractions yield the conclusion of $\mathsf{cut}(R_1, R_2, p)$. □

Using the notion of a cut between two rules of the rule set we can now state the first of our conditions on the rule set.

Definition 2.4.6. A rule $R = \mathcal{P}_R/\Sigma_R \Rightarrow \Pi_R$ *subsumes* a rule $Q = \mathcal{P}_Q/\Sigma_Q \Rightarrow \Pi_Q$ if $\Sigma_R \Rightarrow \Pi_R = \Sigma_Q \Rightarrow \Pi_Q$ and the premisses of Q follow by WCon from the premisses of R, i.e., for every sequent $\Gamma \Rightarrow \Delta$ the sequents in $\mathcal{P}_Q(\Gamma \Rightarrow \Delta)$ are derivable from $\mathcal{P}_R(\Gamma \Rightarrow \Delta)$ using only ConW.

Example 2.4.7. 1. Every rule trivially subsumes itself.

2. The rule $\mathsf{4}_n$ subsumes the rule K_n, since if we have an application of K_n with premiss $A_1, \dots, A_n \Rightarrow B$ and conclusion $\Gamma, \Box A_1, \dots, \Box A_n \Rightarrow \Box B, \Delta$, then from the premiss using only Weakening we can derive the premiss $\Gamma\restriction_{\mathcal{C}_4}, A_1, \dots, A_n \Rightarrow \mathsf{B}, \Delta\restriction_{\mathcal{C}_4}$ of an application of the rule $\mathsf{4}_n$ with the same conclusion.

Definition 2.4.8. A rule set $\mathcal{R}$ is called *principal-cut closed* if for every two rules R_1, R_2 from $\mathcal{R}$ and every formula $\heartsuit\vec{p}$: whenever the rule $\mathsf{cut}(R_1, R_2, \heartsuit\vec{p})$ is defined, then it is subsumed by a rule from $\mathcal{R}$.

Example 2.4.9. 1. The sets $\mathsf{G[cim]}$ of propositional rules are principal-cut closed since all possible principal cuts are subsumed by the identity rule.

2. The rule set $\mathcal{R}_{\mathsf{K}} = \{\mathsf{K}_n \mid n \geq 0\}$ is principal cut closed since for rules

$$\mathsf{K}_n = \{(p_1, \dots, p_n \Rightarrow q; \mathcal{C}_\emptyset)\}/\Box p_1, \dots, \Box p_n \Rightarrow \Box q$$
$$\mathsf{K}_{m+1} = \{(q, q_1, \dots, q_m \Rightarrow r; \mathcal{C}_\emptyset)\}/\Box q, \Box q_1, \dots, \Box q_m \Rightarrow \Box r$$

the rule $\mathsf{cut}(\mathsf{K}_n, \mathsf{K}_{m+1}, \Box q)$ is subsumed by the rule K_{n+m}.

3. Similarly, the rule set $\mathcal{R}_{\mathsf{K4}} := \mathcal{R}_{\mathsf{K}} \cup \{\mathsf{4}_n \mid n \geq 0\}$ is principal-cut closed, since in addition for $m, n \geq 0$ the rule $\mathsf{cut}(\mathsf{4}_m, \mathsf{K}_n, \Box p)$ is subsumed by the rule $\mathsf{4}_{m+n}$ and similarly for cuts between rules $\mathsf{4}_m$ and $\mathsf{4}_n$.

4. The rule set $\mathcal{R}_{\mathsf{K}} \cup \{\mathsf{4}_0\}$ is not principal-cut closed, since e.g. the rule $\mathsf{cut}(\mathsf{K}_2, \mathsf{4}_0, \Box p) = \mathsf{4}_1$ is not subsumed by any rule in the rule set.

5. Finally, the rule sets $\mathcal{R}_{\mathsf{KT}} := \mathcal{R}_{\mathsf{K}} \cup \{\mathsf{T}_n \mid n \geq 1\}$ and $\mathcal{R}_{\mathsf{S4}} := \mathcal{R}_{\mathsf{K4}} \cup \{\mathsf{T}_n \mid n \geq 1\}$ with

$$\mathsf{T}_n := \{(p_1, \dots, p_n \Rightarrow\ ; \mathcal{C}_{\mathsf{id}})\}/\Box p_1, \dots, \Box p_n \Rightarrow$$

are principal-cut closed, but the rule sets $\mathcal{R}_{\mathsf{K}} \cup \{\mathsf{T}_1\}$ and $\mathcal{R}_{\mathsf{K}} \cup \{\mathsf{T}_1, \mathsf{4}_0\}$ are not.

Remark 2.4.10. Under this definition of principal-cut closure e.g. the standard rule sets $\mathcal{R}_{\mathsf{K}} \cup \{\mathsf{T}_1\}$ for modal logic KT and $\mathcal{R}_{\mathsf{K}} \cup \{\mathsf{T}_1, \mathsf{4}_0\}$ for modal logic $\mathsf{S4}$ are as we have seen *not* principal-cut closed. The condition can be weakened to include these rule sets by demanding that the cut between two rules be only *derivable* instead of subsumed by a rule from the rule set (more details will be mentioned in Remark 2.4.18). For a better integration with the remaining conditions we keep to the present condition.

The criterion of principal-cut closure enables us to permute applications of the cut rule above applications of rules with restrictions whenever the cut formula was principal in the two last applied rules. If the cut formula was contextual in at least one of the two last applied rules we would like to be able to permute the application of cut into the premisses of one of the rules. Unfortunately, this might not always be possible, since the context restrictions of this rule might prevent its application below the cut. The following criteria ensure that this is not the case.

Definition 2.4.11. Two restrictions $\mathcal{C}_1 = \langle F_1, F_2 \rangle, \mathcal{C}_2 = \langle G_1, G_2 \rangle$ *overlap* if there are formulae $A_1 \in F_2, A_2 \in G_1$ and substitutions σ_1, σ_2 with $A_1\sigma_1 = A_2\sigma_2$. A rule set $\mathcal{R}$ is

1. *context-cut closed* if whenever $R_0, R_1 \in \mathcal{R}$ and there are context restrictions $\mathcal{C}_0$ of R_0 and $\mathcal{C}_1$ of R_1 which overlap, then there is $i \in \{0, 1\}$ such that all context restrictions of R_i which overlap $\mathcal{C}_{1-i}$ and the principal formulae of R_i satisfy $\mathcal{C}_{1-i}$.

2. *mixed-cut closed* if whenever $R, Q \in \mathcal{R}$ and a principal formula A of R satisfies a context restriction of Q, then all context restrictions of R and all principal formulae of R except for possibly A satisfy all those context restrictions of Q satisfied by A.

Example 2.4.12.

1. Rule sets containing only shallow or one-step rules are trivially context- and mixed-cut closed. Thus all of $\mathsf{G[cim]}, \mathsf{Gc}\mathcal{R}_{\mathsf{K}}, \mathsf{Gc}\mathcal{R}_{\mathsf{KT}}$ are context- and mixed-cut closed.

2. The rule set $\mathcal{R}_{\mathsf{S4}} = \mathcal{R}_{\mathsf{KT}} \cup \{\mathsf{4}_n \mid n \geq 0\}$ is context-cut closed, since for every non-propositional R_0, R_1 with overlapping context restrictions we must have $R_0 = \mathsf{T}_m$ for some m and every formula satisfies the context restriction of T_m.

3. The rule set $\mathcal{R}_{\mathsf{S4}}$ is also mixed-cut closed, since (for the only non-trivial case) the left part of the principal formulae of the rule K_n or $\mathsf{4}_n$ consists only of boxed formulae and thus satisfies the restriction $\mathcal{C}_{\mathsf{4}}$ of rule $\mathsf{4}_m$.

4. Similarly, the rule set $\mathcal{R}_{\mathsf{K4}}$ is context-cut and mixed-cut closed.

5. Every rule set containing the two rules with context restrictions $\langle \emptyset, \{p\} \rangle$ and $\langle \{p\}, \emptyset \rangle$ respectively is not context-cut closed, since these two context relations overlap, but neither satisfies the other. Thus e.g. the system SKL^1 from [Cro01] is not context-cut closed.

6. The well-known rule set $\mathcal{R}_{\mathsf{S5}} := \{R_{\mathsf{45}}, \mathsf{T}_1\}$ is not mixed-cut closed, since the context restriction $\mathcal{C}_{\mathsf{45}}$ of the rule R_{45} is satisfied by the principal formula of T_1, but the context restriction $\mathcal{C}_{\mathsf{id}}$ of T_1 does not satisfy the restriction $\mathcal{C}_{\mathsf{45}}$.

While these criteria might on the surface seem sufficient to ensure Cut Elimination, we need another condition on the rule set, which enables us to permute certain applications of the contraction rule above applications of rules with restrictions. This is necessary to deal with the problematic case that the cut formula is the principal formula in applications of the contraction rule in both premisses of the cut, and that on each side at least two instances of the cut formula have been introduced by a rule from the rule set. On the other hand, this is not too much of a restriction, since we will also need it later on in Section 2.6 to show that Contraction is admissible in the system with slightly modified rules. Similarly to the case of cut we first lift the notion of contraction from sequents to sequent rules.

Definition 2.4.13. If $\mathcal{P}$ is a set of premisses with restrictions and $\vec{p} = (p_1, \dots, p_n)$ and $\vec{q} = (q_1, \dots, q_n)$ are n-tuples of variables, then $\mathcal{P}[\vec{q} \leftarrow \vec{p}]$ is the result of replacing every occurrence of q_i in a sequent occurring in a premiss in $\mathcal{P}$ by p_i for all $i = 1, \dots, n$ and contracting duplicate instances of $p_1, \dots, p_n$. Let $R = \mathcal{P}/\Sigma, \heartsuit\vec{p}, \heartsuit\vec{q} \Rightarrow \Pi$ be a rule. The *left contraction of R on $\heartsuit\vec{p}$ and $\heartsuit\vec{q}$* is the rule $\mathsf{ConL}(R, \heartsuit\vec{p}, \heartsuit\vec{q}) = \mathcal{P}[\vec{q} \leftarrow \vec{p}]/\Sigma, \heartsuit\vec{p} \Rightarrow \Pi$. The *right contraction* $\mathsf{ConR}(R, \heartsuit\vec{p}, \heartsuit\vec{q})$ is defined dually. A rule set $\mathcal{R}$ is *left-contraction closed* (resp. *right-contraction closed*), if for every rule R from $\mathcal{R}$ applications of the rules $\mathsf{ConL}(R, \heartsuit\vec{p}, \heartsuit\vec{q})$ (resp. $\mathsf{ConR}(R, \heartsuit\vec{p}, \heartsuit\vec{q})$) can be simulated by applications of Weakening and Contraction followed by at most one application of a rule R' from $\mathcal{R}$ and Weakening. A rule set is *contraction closed* if it is left- and right-contraction closed and *saturated* if it is contraction, principal-cut, context-cut, and mixed-cut closed.

Example 2.4.14.
1. Every rule set $\mathcal{R}$ in which the principal formulae of every rule have the form $\Gamma \Rightarrow \delta$ (resp. $\gamma \Rightarrow \Delta$) is trivially right- (resp. left-)contraction closed. Thus in the asymmetric framework every rule set is trivially right-contraction closed (but not necessarily contraction closed).

2. For $\mathcal{P} = \{(q_1, \dots, q_n, q_{n+1} \Rightarrow r; \mathcal{C}_\emptyset)\}$ we have $\mathcal{P}[q_{n+1} \leftarrow q_n] = \{(q_1, \dots, q_n \Rightarrow r; \mathcal{C}_\emptyset)\}$. Thus for the rule $\mathsf{K}_{n+1} = \{(q_1, \dots, q_n, q_{n+1} \Rightarrow r; \mathcal{C}_\emptyset)\}/\Box q_1, \dots, \Box q_n, \Box q_{n+1} \Rightarrow \Box r$ we have $\mathsf{ConL}(\mathsf{K}_n, \Box q_n, \Box q_{n+1}) = \{(q_1, \dots, q_n \Rightarrow r; \mathcal{C}_\emptyset)\}/\Box q_1, \dots, \Box q_n \Rightarrow \Box r$. Thus the set $\mathcal{R}_{\mathsf{K}}$ is left-contraction closed. It is also trivially right-contraction closed and thus contraction closed. Together with Examples 2.4.9 and 2.4.12 this shows that $\mathcal{R}_{\mathsf{K}}$ is saturated.

3. Similarly, the rule sets $\mathcal{R}_{\mathsf{KT}}, \mathcal{R}_{\mathsf{K4}}$ and $\mathcal{R}_{\mathsf{S4}}$ are contraction closed and thus with Examples 2.4.9 and 2.4.12 saturated.

4. The rule set $\{R_{\mathsf{K}_2}\}$ is trivially right-contraction closed, but not left-contraction closed.

5. Since the principal formulae of the propositional rules contain at most one formula on the left resp. right hand side, the rule sets $\mathsf{G[cim]}$ are trivially left- and right-contraction closed and thus contraction closed and with Examples 2.4.9 and 2.4.12 saturated.

The benefit of demanding that the rule sets are left- or right-contraction closed lies in the fact that if a rule set is right-contraction closed, then we can permute contractions on the right hand side of the principal formulae of a rule into the premisses of this rule (and dually for left-contraction closed rule sets). This gives us the following lemma.

Lemma 2.4.15. *Let $\mathcal{R}$ be a left-contraction closed (resp. right-contraction closed) set of rules. If a sequent is derivable in $\mathcal{R}\mathsf{Con[CutW]}$, then there is a derivation of it in $\mathcal{R}\mathsf{Con[CutW]}$, in which in every application of a rule from $\mathcal{R}$ every formula occurs at most once on the left (resp. right) hand side of the principal formulae of this application.*

Proof. We show the statement for left-contraction closed rule sets. The proof for right-contraction closed rule sets is analogous.

We first show that whenever we have a derivation of a sequent $\Gamma, A^k \Rightarrow \Delta$ which ends in an application of a rule R from $\mathcal{R}$ and in which in every application of a rule from $\mathcal{R}$ above this application of R every formula occurs at most once on the left hand side of the principal formulae, we can transform this into a derivation where in *every* application of a rule from $\mathcal{R}$ every formula occurs at most once on the left hand side of the principal formulae. The statement of the Lemma then follows by repeatedly eliminating topmost applications of contraction in the principal formulae. The proof is by induction on $\mathsf{rk}(A)$.

If $\mathsf{rk}(A) = 0$, then due to the format of rules with restrictions the formula A cannot occur on the left hand side of the principal formulae of an application of a rule from $\mathcal{R}$ and thus the derivation already has the desired form.

So suppose that $\mathsf{rk}(A) = n + 1$. Then the derivation ends as follows:

$$\dfrac{\begin{array}{c}\mathcal{D}_1\\ \vdots\\ \Gamma_1 \Rightarrow \Delta_1\end{array} \qquad \dots \qquad \begin{array}{c}\mathcal{D}_n\\ \vdots\\ \Gamma_n \Rightarrow \Delta_n\end{array}}{\Sigma, A^k \Rightarrow \Pi}\ R$$

where all k displayed instances of the formula A are principal in the application of the rule R. Possible additional instances of A in the context are not problematic. Since the rule set is left-contraction closed we can turn this into

$$\begin{array}{c}\dfrac{\begin{array}{c}\mathcal{D}_1\\ \vdots\\ \Gamma_1 \Rightarrow \Delta_1\\ \hline\hline \Gamma_1' \Rightarrow \Delta_1'\end{array}\ \mathsf{ConW} \qquad \dots \qquad \begin{array}{c}\mathcal{D}_1\\ \vdots\\ \Gamma_1 \Rightarrow \Delta_1\\ \hline\hline \Gamma_1' \Rightarrow \Delta_1'\end{array}\ \mathsf{ConW}}{\Sigma', A^{k-i} \Rightarrow \Pi'}\ Q\\ \hline\hline \Sigma, A^k \Rightarrow \Pi\end{array}\ \mathsf{W}$$

where Q is a rule in $\mathcal{R}$ and where again all $k-i$ displayed instances of A in the conclusion of the application of Q are principal. Moreover, it can be seen that the contractions in the premisses must be on proper subformulae of A, and thus are on formulae of rank at most n. Any newly introduced contractions of formulae occurring on the left hand side of the principal formulae of an application of a rule in this derivation therefore must be on formulae of rank at most n and are eliminated using the induction hypothesis. Repeating this proces we eliminate the remaining duplicates of the formula A in the conclusion of the application of Q. Finally, if Weakening is not in the rule set, we use admissibility of Weakening (Lemma 2.3.12) to eliminate applications of W. Note that this introduces additional instances of formulae only in the context and not in the principal formulae of applications of rules. $\square$

Using this lemma if we have a right-contraction closed rule set, then we may assume w.l.o.g. that in a derivation the right hand sides of the principal formulae of applications of rules contain at most one instance of every formula. Thus the above mentioned problematic case in the proof of cut elimination does not occur and we can build on the techniques of [vP01] and [PS08] to show cut elimination.

Theorem 2.4.16 (Cut Elimination)**.** *Let $\mathcal{R}$ be a cut closed set of rules with restrictions. If $\mathcal{R}$ is left-contraction closed or right-contraction closed, then the cut rule can be eliminated, i.e. for every sequent $\Gamma \Rightarrow \Delta$ we have $\vdash_{\mathcal{R}\mathsf{Cut}[\mathsf{Con}]} \Gamma \Rightarrow \Delta$ if and only if $\vdash_{\mathcal{R}[\mathsf{Con}]} \Gamma \Rightarrow \Delta$.*

Proof. We show the theorem for right-contraction closed rule sets. The proof for left-contraction closed rule sets is similar. For an application

$$\dfrac{\begin{array}{c}\mathcal{D}_1\\ \vdots\\ \Gamma \Rightarrow \Delta, A\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ \vdots\\ A, \Sigma \Rightarrow \Pi\end{array}}{\Gamma, \Sigma \Rightarrow \Delta, \Pi}\ \mathsf{Cut}$$

of the cut rule we call $\Gamma \Rightarrow \Delta, A$ the *left premiss* and $A, \Sigma \Rightarrow \Pi$ the *right premiss* of this application. Furthermore, we say that the *rank* of this application is the rank of the cut formula A, and its *height* is the sum of the depths of the two derivations $\mathcal{D}_1$ and $\mathcal{D}_2$ of its left resp. right premisses. In the context of this proof we say that a derivation $\mathcal{D}$ has *property* (P) if in every application of a rule from $\mathcal{R}$ every formula occurs at most once on the right hand side of the principal formulae of this application. Given a derivation $\mathcal{D}$ with property (P) as usual we transform topmost applications of Cut in it into (possibly several) applications of Cut with lower rank or equal rank and lower height. Then by a double induction on the rank and the height of the applications of Cut we eliminate all applications of Cut in $\mathcal{D}$, where after eliminating each application we appeal to Lemma 2.4.15 to ensure that the resulting derivation again has property (P). While Lemma 2.4.15 does not guarantee that the depth of the derivation is preserved, this is not a problem, since cuts with lower height than the original cut will only be followed by cuts with lower rank.

So suppose we have a topmost application of Cut. Then this has the form

$$\dfrac{\dfrac{\dfrac{\begin{array}{c}\mathcal{E}_1\\ \vdots \\ \Gamma_1 \Rightarrow \Delta_1\end{array} \quad \dots \quad \begin{array}{c}\mathcal{E}_i\\ \vdots \\ \Gamma_i \Rightarrow \Delta_i\end{array}}{\Gamma' \Rightarrow \Delta', A^n}\,R}{\Gamma \Rightarrow \Delta, A}\,\mathsf{Con} \qquad \dfrac{\dfrac{\begin{array}{c}\mathcal{F}_1\\ \vdots \\ \Sigma_1 \Rightarrow \Pi_1\end{array} \quad \dots \quad \begin{array}{c}\mathcal{F}_j\\ \vdots \\ \Sigma_j \Rightarrow \Pi_j\end{array}}{A^k, \Sigma' \Rightarrow \Pi'}\,Q}{A, \Sigma \Rightarrow \Pi}\,\mathsf{Con}}{\Gamma, \Sigma \Rightarrow \Delta, \Pi}\,\mathsf{Cut}$$

where R and Q are rules in $\mathcal{R}$ or applications of the axiom rule. In the latter case we take the number of premisses to be 0. In a first step we permute contractions so that all contractions of the cut formula occur just below the applications of R and Q. This gives

$$\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{c}\mathcal{E}_1\\ \vdots \\ \Gamma_1 \Rightarrow \Delta_1\end{array} \quad \dots \quad \begin{array}{c}\mathcal{E}_i\\ \vdots \\ \Gamma_i \Rightarrow \Delta_i\end{array}}{\Gamma' \Rightarrow \Delta', A^n}\,R}{\Gamma' \Rightarrow \Delta', A}\,\mathsf{ConR}}{\Gamma \Rightarrow \Delta, A}\,\mathsf{Con} \qquad \dfrac{\dfrac{\dfrac{\begin{array}{c}\mathcal{F}_1\\ \vdots \\ \Sigma_1 \Rightarrow \Pi_1\end{array} \quad \dots \quad \begin{array}{c}\mathcal{F}_j\\ \vdots \\ \Sigma_j \Rightarrow \Pi_j\end{array}}{A^k, \Sigma' \Rightarrow \Pi'}\,Q}{A, \Sigma' \Rightarrow \Pi'}\,\mathsf{ConL}}{A, \Sigma \Rightarrow \Pi}\,\mathsf{Con}}{\Gamma, \Sigma \Rightarrow \Delta, \Pi}\,\mathsf{Cut}$$

Note that this does not change the height or the rank of the cut.

Now suppose that the last applied rule in at least one of the two premisses of the cut was a contraction rule where the cut formula is not principal. We show the transformation for the rule ConL, the case of ConR is analogous. We have one of the two following situations:

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_1'\\ \vdots \\ \Gamma, B, B \Rightarrow \Delta, A\end{array}}{\Gamma, B \Rightarrow \Delta, A}\,\mathsf{ConL} \qquad \begin{array}{c}\mathcal{D}_2\\ \vdots \\ A, \Sigma \Rightarrow \Pi\end{array}}{\Gamma, B, \Sigma \Rightarrow \Delta, \Pi}\,\mathsf{Cut} \qquad \dfrac{\begin{array}{c}\mathcal{D}_1'\\ \vdots \\ \Gamma \Rightarrow \Delta, A\end{array} \qquad \dfrac{\begin{array}{c}\mathcal{D}_2'\\ \vdots \\ A, B, B, \Sigma \Rightarrow \Pi\end{array}}{A, B, \Sigma \Rightarrow \Pi}\,\mathsf{ConL}}{\Gamma, \Sigma, B \Rightarrow \Delta, \Pi}\,\mathsf{Cut}$$

By permuting the application of Cut above the applications of ConL these are transformed into the derivations

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_1'\\ \vdots \\ \Gamma, B, B \Rightarrow \Delta, A\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ \vdots \\ A, \Sigma \Rightarrow \Pi\end{array}}{\Gamma, B, B, \Sigma \Rightarrow \Delta, \Pi}\,\mathsf{Cut}}{\Gamma, B, \Sigma \Rightarrow \Delta, \Pi}\,\mathsf{ConL} \qquad \dfrac{\dfrac{\begin{array}{c}\mathcal{D}_1\\ \vdots \\ \Gamma \Rightarrow \Delta, A\end{array} \qquad \begin{array}{c}\mathcal{D}_2'\\ \vdots \\ A, \Sigma, B, B \Rightarrow \Pi\end{array}}{\Gamma, \Sigma, B, B \Rightarrow \Delta, \Pi}\,\mathsf{Cut}}{\Gamma, \Sigma, B \Rightarrow \Delta, \Pi}\,\mathsf{ConL}$$

where the application of Cut has the same rank and lower height than the original cut.

Otherwise all contractions were on the cut formula and we have the situation

$$\dfrac{\dfrac{\dfrac{\begin{matrix}\mathcal{E}_1 & & \mathcal{E}_i\\ \vdots & & \vdots \\ \Gamma_1 \Rightarrow \Delta_1 & \dots & \Gamma_i \Rightarrow \Delta_i\end{matrix}}{\Gamma \Rightarrow \Delta, A^n}\,R}{\Gamma \Rightarrow \Delta, A}\,\mathsf{ConR} \qquad \dfrac{\dfrac{\begin{matrix}\mathcal{F}_1 & & \mathcal{F}_j\\ \vdots & & \vdots \\ \Sigma_1 \Rightarrow \Pi_1 & \dots & \Sigma_j \Rightarrow \Pi_j\end{matrix}}{A^k, \Sigma \Rightarrow \Pi}\,Q}{A, \Sigma \Rightarrow \Pi}\,\mathsf{ConL}}{\Gamma, \Sigma \Rightarrow \Delta, \Pi}\,\mathsf{Cut}$$

where again R and Q are rules from $\mathcal{R}$ or the axiom rule.

If at least one of R and Q was the axiom rule, we have the following cases.

1. $R = \mathcal{A}$ and A is contextual in R. Then we have (abbreviating the derivation of the right premiss $A, \Sigma \Rightarrow \Pi$ with $\mathcal{F}$):

$$\dfrac{\dfrac{\dfrac{}{\Gamma', B \Rightarrow \Delta', B, A^m}\,\mathcal{A}}{\Gamma', B \Rightarrow \Delta', B, A}\,\mathsf{ConR} \qquad \begin{matrix}\mathcal{F}\\ \vdots \\ A, \Sigma \Rightarrow \Pi\end{matrix}}{\Gamma', B, \Sigma \Rightarrow \Delta', B, \Pi}\,\mathsf{Cut}$$

which is transformed into the cut-free derivation

$$\dfrac{}{\Gamma', B, \Sigma \Rightarrow \Delta', B, \Pi}\,\mathcal{A}$$

2. $R = \mathcal{A}$ and A is principal in R. Then we have (again abbreviating the derivation of the right premiss of the cut):

$$\dfrac{\dfrac{\dfrac{}{\Gamma', A \Rightarrow \Delta, A^m}\,\mathcal{A}}{\Gamma', A \Rightarrow \Delta, A}\,\mathsf{ConR} \qquad \begin{matrix}\mathcal{F}\\ \vdots \\ A, \Sigma \Rightarrow \Pi\end{matrix}}{\Gamma', A, \Sigma \Rightarrow \Delta, \Pi}\,\mathsf{Cut}$$

and the application of Cut can be omitted using admissibility of Weakening.

3. $R \neq \mathcal{A}$ and $Q = \mathcal{A}$ and A is contextual in Q. Dual to Case 1.

4. $R \neq \mathcal{A}$ and $Q = \mathcal{A}$ and A is principal in Q. Dual to Case 2.

Now suppose that neither of R, Q is the axiom rule. Then instances of the cut formula A might be introduced by the applications of R resp. Q or be part of the context. For the sake of presentation we assume that every premiss of R and Q carries over formulae of the form A. The treatment of rules involving premisses not carrying over formulae of this form is similar.

Thus for some $n, m, k, \ell \geq 0$ we have:

$$
\dfrac{
\dfrac{
\dfrac{
\begin{array}{c}\mathcal{E}_1\\ \vdots\\ \Gamma_1 \Rightarrow \Delta_1, A^m\end{array}
\quad \dots \quad
\begin{array}{c}\mathcal{E}_i\\ \vdots\\ \Gamma_i \Rightarrow \Delta_i, A^m\end{array}
}{\Gamma \Rightarrow \Delta, A^n, A^m}\, R
}{\Gamma \Rightarrow \Delta, A}\, \mathsf{ConR}
\qquad
\dfrac{
\dfrac{
\begin{array}{c}\mathcal{F}_1\\ \vdots\\ A^k, \Sigma_1 \Rightarrow \Pi_1\end{array}
\quad \dots \quad
\begin{array}{c}\mathcal{F}_j\\ \vdots\\ A^k, \Sigma_j \Rightarrow \Pi_j\end{array}
}{A^k, A^\ell, \Sigma \Rightarrow \Pi}\, Q
}{A, \Sigma \Rightarrow \Pi}\, \mathsf{ConL}
}{\Gamma, \Sigma \Rightarrow \Delta, \Pi}\, \mathsf{Cut}
$$

Note that since the derivation has property (P), the formula A occurs at most once in the principal formulae of the application of R, and thus we have $n \in \{0, 1\}$. This derivation is now transformed into a derivation of $\Gamma, \Sigma \Rightarrow \Delta, \Pi$ using several cuts of the same rank and lower height or of lower rank as follows.

Consider the case that $n + \ell > 0$, that is that not all instances of A are contextual. First if $m > 0$, then we eliminate the instances of the cut formula A in the premisses of the application of R using the derivation

$$
\dfrac{
\dfrac{
\dfrac{
\begin{array}{c}\mathcal{E}_1\\ \vdots\\ \Gamma_1 \Rightarrow \Delta_1, A^m\end{array}
}{\Gamma_1 \Rightarrow \Delta_1, A}\, \mathsf{Con}
\quad
\dfrac{
\dfrac{
\begin{array}{c}\mathcal{F}_1\\ \vdots\\ A^k, \Sigma_1 \Rightarrow \Pi_1\end{array}
\quad \dots \quad
\begin{array}{c}\mathcal{F}_j\\ \vdots\\ A^k, \Sigma_j \Rightarrow \Pi_j\end{array}
}{A^k, A^\ell, \Sigma \Rightarrow \Pi}\, Q
}{A, \Sigma \Rightarrow \Pi}\, \mathsf{Con}
}{\Gamma_1, \Sigma \Rightarrow \Delta_1, \Pi}\, \mathsf{Cut}
\quad \dots \quad
\dfrac{\dots \qquad \dots}{\Gamma_i, \Sigma \Rightarrow \Delta_i, \Pi}\, \mathsf{Cut}
}{\Gamma, \Sigma \Rightarrow \Delta, A^n, \Pi}\, R
$$

The cut can be permuted into the premisses of R since the rule set is mixed-cut closed and thus the additional formulae $\Sigma \Rightarrow \Pi$ in the premisses satisfy all context restrictions occurring in R. All the newly introduced cuts have the same rank and lower height than the original cut. If $n = 0$, that is if all instances of A in the application of R were contextual, we are done. Otherwise we have $n = 1$. For $s \in \{1, \dots, i\}$ let us write $\mathcal{G}_s$ for the subderivation of the premiss $\Gamma_s, \Sigma \Rightarrow \Delta_s, \Pi$ of R. In case $m = 0$ we set $\mathcal{G}_s = \mathcal{E}_s$. Similar to above, if $k > 0$ we eliminate the instances of the cut formula in the premisses of the application of Q using

$$
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\begin{array}{c}\mathcal{E}_1\\ \vdots\\ \Gamma_1 \Rightarrow \Delta_1, A^m\end{array}
\quad \dots \quad
\begin{array}{c}\mathcal{E}_n\\ \vdots\\ \Gamma_n \Rightarrow \Delta_n, A^m\end{array}
}{\Gamma \Rightarrow \Delta, A^m}\, R
}{\Gamma \Rightarrow \Delta, A}\, \mathsf{Con}
\quad
\dfrac{
\begin{array}{c}\mathcal{F}_1\\ \vdots\\ A^k, \Sigma_1 \Rightarrow \Pi_1\end{array}
}{A, \Sigma_1 \Rightarrow \Pi_1}\, \mathsf{Con}
}{\Gamma, \Sigma_1 \Rightarrow \Delta, \Pi_1}\, \mathsf{Cut}
\quad \dots \quad
\dfrac{\dots \qquad \dots}{\Gamma, \Sigma_j \Rightarrow \Delta, \Pi_j}\, \mathsf{Cut}
}{\Gamma, A^\ell, \Sigma \Rightarrow \Delta, \Pi}\, Q
$$

Again the cuts can be permuted into the premisses of the application of Q since the rule set is mixed-cut closed, and all the newly introduced cuts have the same rank and lower height than the original cut. If $\ell = 0$, then we are done. Otherwise for $t \in \{1, \dots, j\}$ we write $\mathcal{H}_t$ for the derivation of the premiss $\Gamma, \Sigma_t \Rightarrow \Delta, \Pi_t$ of this application of Q. If $k = 0$ we set $\mathcal{H}_t = \mathcal{F}_t$. Now we piece these derivations together using cuts to get

$$
\begin{array}{c}
\begin{array}{cc}
\begin{array}{c}\mathcal{G}_1 \qquad \mathcal{G}_n\\ \vdots \quad \dots \quad \vdots\\ \hline \Gamma, \Sigma \Rightarrow \Delta, A, \Pi\end{array} R
&
\begin{array}{c}
\begin{array}{cc}
\begin{array}{c}\mathcal{G}_1 \qquad \mathcal{G}_n\\ \vdots \quad \dots \quad \vdots\\ \hline \Gamma, \Sigma \Rightarrow \Delta, A, \Pi\end{array} R &
\begin{array}{c}\mathcal{H}_1 \qquad \mathcal{H}_j\\ \vdots \quad \dots \quad \vdots\\ \hline \Gamma, A^\ell, \Sigma \Rightarrow \Delta, \Pi\end{array} Q
\end{array}\\ \hline
\Gamma^2, A^{\ell-1}, \Sigma^2 \Rightarrow \Delta^2, \Pi^2 \quad \mathsf{Cut}\\ \vdots\\ \Gamma^\ell, A, \Sigma^\ell \Rightarrow \Delta^\ell, \Pi^\ell
\end{array}
\end{array}\\ \hline
\Gamma^{\ell+1}, \Sigma^{\ell+1} \Rightarrow \Delta^{\ell+1}, \Pi^{\ell+1} \quad \mathsf{Cut}\\ \hline\hline
\Gamma, \Sigma \Rightarrow \Delta, \Pi \quad \mathsf{Con}
\end{array}
$$

The newly introduced cuts still have the same rank but possibly greater height than the original cut. But since the cut formula A is principal in the topmost application of R and the application of Q, and since the rule set is principal-cut closed, the rule $\mathsf{cut}(R, Q, A)$ is subsumed by a rule $S_{\mathsf{cut}(R,Q,A)}$ in $\mathcal{R}$ and we can replace the topmost cut by cuts on proper subformulae of A followed by Weakenings and Contractions and an application of the rule $S_{\mathsf{cut}(R,Q,A)}$. This gives

$$
\begin{array}{c}
\begin{array}{cc}
\begin{array}{c}\mathcal{G}_1 \qquad \mathcal{G}_n\\ \vdots \quad \dots \quad \vdots\\ \hline \Gamma, \Sigma \Rightarrow \Delta, A, \Pi\end{array} R
&
\begin{array}{c}
\begin{array}{ccc}
\begin{array}{c}\mathcal{G}_1 \qquad \mathcal{G}_n \quad \mathcal{H}_1 \qquad \mathcal{H}_j\\ \vdots \quad \dots \quad \vdots \quad\; \vdots \quad \dots \quad \vdots\\ \hline\hline \Theta_1 \Rightarrow \Xi_1 \quad \mathsf{Cut}\\ \hline\hline \Theta'_1 \Rightarrow \Xi'_1 \quad \mathsf{ConW}\end{array} & \dots &
\begin{array}{c}\mathcal{G}_1 \qquad \mathcal{G}_n \quad \mathcal{H}_1 \qquad \mathcal{H}_j\\ \vdots \quad \dots \quad \vdots \quad\; \vdots \quad \dots \quad \vdots\\ \hline\hline \Theta_r \Rightarrow \Xi_r \quad \mathsf{Cut}\\ \hline\hline \Theta'_r \Rightarrow \Xi'_r \quad \mathsf{ConW}\end{array}
\end{array}\\ \hline
\Gamma^2, A^{\ell-1}, \Sigma^2 \Rightarrow \Delta^2, \Pi^2 \quad S_{\mathsf{cut}(R,Q,A)}\\ \vdots\\ \Gamma^\ell, A, \Sigma^\ell \Rightarrow \Delta^\ell, \Pi^\ell
\end{array}
\end{array}\\ \hline
\Gamma^{\ell+1}, \Sigma^{\ell+1} \Rightarrow \Delta^{\ell+1}, \Pi^{\ell+1} \quad \mathsf{Cut}\\ \hline\hline
\Gamma, \Sigma \Rightarrow \Delta, \Pi \quad \mathsf{Con}
\end{array}
$$

Note that since the remaining $\ell - 1$ occurrences of A in the conclusion of $S_{\mathsf{cut}(R,Q,A)}$ were principal in Q, they are principal in $\mathsf{cut}(R, Q, A)$, and thus also principal in $S_{\mathsf{cut}(R,Q,A)}$. Moreover, if $S_{\mathsf{cut}(R,Q,A)}$ was the identity rule R_{id}, then we must have $\ell = 1$ and this was the only remaining cut on A. Otherwise, continuing like this we replace all the remaining cuts on A by cuts on proper subformulae of A and applications of rules from the rule set. Since this does not change the cuts in the derivations $\mathcal{G}_s$ and $\mathcal{H}_t$, all the cuts in the resulting derivation of $\Gamma, \Sigma \Rightarrow \Delta, \Pi$ thus have either the same rank as the original cut and lower height or lower rank.

In the remaining case we have $n = \ell = 0$, and in the original derivation all instances of A are contextual in the applications of R and Q. Then if $m = 0$ or $k = 0$ we get the sequent $\Gamma, \Sigma \Rightarrow \Delta, \Pi$ from one of the conclusions of R, Q by admissibility of weakening, thus eliminating the application of Cut. On the other hand, if $m \neq 0$ and $k \neq 0$, then since the rule set is context-cut closed we know that we can permute the application of Cut into the premisses of the application of R or the application of Q. The derivation then is transformed as in the first steps of the previous case. □

As we have seen in Example 2.4.14 the standard systems $\mathsf{G[cim]}$ of propositional rules and the systems $\mathsf{Gc}\mathcal{R}_{\mathsf{K}}, \mathsf{Gc}\mathcal{R}_{\mathsf{K4}}, \mathsf{Gc}\mathcal{R}_{\mathsf{KT}}$ and $\mathsf{Gc}\mathcal{R}_{\mathsf{S4}}$ are saturated sets of rules with restrictions in the sense of Definition 2.4.13. Thus we obtain cut elimination for these systems as a corollary from the previous theorem.

Corollary 2.4.17. *The sequent calculi given by the rules* $\mathsf{G[cim]}$ *as well as* $\mathsf{Gc}\mathcal{R}_{\mathsf{K}}, \mathsf{Gc}\mathcal{R}_{\mathsf{K4}}, \mathsf{Gc}\mathcal{R}_{\mathsf{KT}}$ *and* $\mathsf{Gc}\mathcal{R}_{\mathsf{S4}}$ *have cut elimination.* □

Remark 2.4.18. It is also possible to show a slightly different version of the generic cut elimination theorem based on a more lenient definition of principal-cut closed rule sets. Let us call a rule set $\mathcal{R}$ *principal-cut deriving* if for every two rules R_1, R_2 from $\mathcal{R}$ the rule $\mathsf{cut}(R_1, R_2, \heartsuit\vec{p})$ is derivable in $\mathcal{R}\mathsf{ConW}$. Then we can show the following analogue of Theorem 2.4.16, where we strengthen the condition of left- or right-contraction closure of the rule set $\mathcal{R}$ in to full contraction closure:

Let $\mathcal{R}$ be a principal-cut deriving, mixed-cut closed, context-cut closed and contraction closed set of rules with restrictions. Then $\mathcal{R}\mathsf{Con}$ has cut elimination, i.e. for every sequent $\Gamma \Rightarrow \Delta$ we have $\vdash_{\mathcal{R}\mathsf{ConCut}} \Gamma \Rightarrow \Delta$ iff $\vdash_{\mathcal{R}\mathsf{Con}} \Gamma \Rightarrow \Delta$.

The proof uses Lemma 2.4.15 to ensure that no formula occurs more than once in the principal formulae of each application of a rule in a derivation of $\Gamma \Rightarrow \Delta$, and then proceeds in the spirit of [Gen34] to eliminate applications of the *multicut rule*

$$\frac{\Gamma \Rightarrow \Delta, A^m \qquad A^n, \Sigma \Rightarrow \Pi}{\Gamma, \Sigma \Rightarrow \Delta, \Pi}\ \mathsf{Mcut}$$

via double induction on the rank of the cut formula and the sum of the depths of the derivations of the two premisses of the multicut. The remainder of the proof is essentially the same as that given for Theorem 2.4.16, with the difference that contractions of the cut formula are absorbed into the application of multicut. In case the last applied rules in the derivations of both premisses of the multicut were rules from $\mathcal{R}$ we again use mixed-cut closure and context-cut closure of the rule set and the same technique to first eliminate all duplicates of the cut formula in the contexts. Then since by the lemma the cut formula occurs only once in the principal formulae of both rules we now only need to eliminate one more cut. We do this by first replacing it with the cut between the two rules, and then replacing this rule by a derivation in $\mathcal{R}\mathsf{ConW}$ using the fact that by principal-cut derivability it is derivable in this rule set. The newly introduced cuts now have smaller rank and are eliminated as above.

Obviously the condition of being principal-cut deriving is weaker than that of being principal-cut closed. In particular the standard rule sets $\mathcal{R}_{\mathsf{K}} \cup \{\mathsf{T}_1\}$ and $\mathcal{R}_{\mathsf{K}} \cup \{\mathsf{T}_1, \mathsf{4}_0\}$ for modal logics

KT and $\mathsf{S4}$ are principal-cut deriving but not principal-cut closed. But on the other hand for these rule sets we can also show cut elimination by taking a derivation in e.g. $\mathcal{R}_{\mathsf{K}} \cup \{\mathsf{T}_1\}$ to be a derivation in $\mathcal{R}_{\mathsf{KT}}$, eliminating the cuts using Theorem 2.4.16, and replacing in the resulting cut free derivation all the rules T_n by their derivations in $\mathcal{R}_{\mathsf{K}} \cup \{\mathsf{T}_1\}$. For this reason and since it better integrates with the rest of the conditions we use the condition of being principal-cut closed in the following.

2.5 Cut Elimination and Invertibility

While the proof of cut elimination via saturation of the rule set has the advantage of being very general and in particular of applying to the propositional rules themselves as well there is another possibility for showing cut elimination. Instead of using mixed-cut closure to push cuts on a propositional formula which is principal in one rule and contextual in the other into the premisses of the second rule, we might use invertibility of the propositional rules to replace this cut by cuts of lower rank. Let us make this precise. Since in general in asymmetric sequent systems not all the propositional rules are invertible in this section we consider only the symmetric framework.

Definition 2.5.1. Let $\mathcal{R}$ be a set of rules with context restrictions. A rule $R = \mathcal{P}/\Sigma \Rightarrow \Pi$ is called *invertible in* $\mathcal{R}[\mathsf{Con}]$ if for every context $\Theta \Rightarrow \Xi$ and substitution σ we have: whenever $\vdash_{\mathcal{R}[\mathsf{Con}]} \Theta, \Sigma\sigma \Rightarrow \Pi\sigma, \Xi$, then for every premiss $(\Gamma \Rightarrow \Delta; \mathcal{C})$ from $\mathcal{P}$ we have $\vdash_{\mathcal{R}[\mathsf{Con}]} \Theta \restriction_{\mathcal{C}}, \Gamma\sigma \Rightarrow \Delta\sigma, \Xi \restriction_{\mathcal{C}}$.

Example 2.5.2. Let $\mathcal{R}$ be the rule set Gc. Then as is well-known e.g. the rule $\rightarrow_{\mathsf{L}} = \{(p \Rightarrow ; \mathcal{C}_{\mathsf{id}}), (\Rightarrow q; \mathcal{C}_{\mathsf{id}})\}/p \to q \Rightarrow$ is invertible in Gc, since whenever for a context $\Theta \Rightarrow \Xi$ and a substitution σ the sequent $\Theta, p\sigma \to q\sigma \Rightarrow \Xi$ is derivable in Gc, then so are the sequents $\Theta, p\sigma \Rightarrow \Xi$ and $\Theta \Rightarrow q\sigma, \Xi$.

The standard way to show that the propositional rules are invertible in Gc is to use a permutation of rules argument similar to the condition of mixed-cut closure, but permuting applications of the propositional rules *below* applications of the rules in $\mathcal{R}$. Similarly to the case for cut elimination we can generalise this method and distill the following sufficient criterion for invertibility of the propositional rules from it.

Definition 2.5.3 (c)**.** A rule set $\mathcal{R}$ is Gc-*inverting* if for every restriction $\langle F_0; F_1\rangle$ of a rule in $\mathcal{R}$ and $i \in \{0, 1\}$ we have: whenever $A \circ B \in F_i$ with $\circ \in \{\wedge, \vee\}$, then also $A, B \in F_i$ and whenever $A \to B \in F_i$, then also $A \in F_{1-i}$ and $B \in F_i$.

Example 2.5.4 (c)**.**

1. Since the context restrictions $\mathcal{C}_{\mathsf{id}}$ and $\mathcal{C}_{\emptyset}$ satisfy the requirements given in Definition 2.5.3 it is clear that every set of shallow rules is Gc-inverting.

2. Since the propositional connectives do not occur in the context restriction $\mathcal{C}_{\mathsf{4}} = \langle \{\Box p\}, \emptyset \rangle$, this restriction also satisfies the requirements from Definition 2.5.3 and together with the above argument for the rules K_n and T_n we have that the rule sets $\mathcal{R}_{\mathsf{K4}}$ and $\mathcal{R}_{\mathsf{S4}}$ are Gc-inverting.

And indeed, if a rule set $\mathcal{R}$ is Gc-inverting we can show invertibility of the rules in Gc via a permutation-of-rules argument.

Lemma 2.5.5 (Inversion Lemma)(c)**.** *Let $\mathcal{R}$ be a Gc-inverting set of rules. Then the propositional rules are invertible in $\mathsf{Gc}\mathcal{R}[\mathsf{Con}]$.*

Proof. Since our rule set might include the rule Con we need to adapt the standard proof a bit. Similar to Gentzen's original proof of cut elimination [Gen34] instead of showing the result for the set $\mathsf{Gc}\mathcal{R}[\mathsf{Con}]$ we show it for the system $\mathsf{G'c}\mathcal{R}[\mathsf{Mcon}]$. Here Mcon is the rule

$$\frac{\Gamma, \Gamma, \Sigma \Rightarrow \Delta, \Delta, \Pi}{\Gamma, \Sigma \Rightarrow \Delta, \Pi}\ \mathsf{Mcon}$$

which allows to contract several different formulae at the same time, and $\mathsf{G'c}$ has axioms $\overline{\Gamma, p \Rightarrow p, \Delta}$ which allow only propositional variables as the principal formulae instead of the general axioms $\overline{\Gamma, A \Rightarrow A, \Delta}$. Obviously the rule Mcon can be simulated by multiple applications of the rules $\mathsf{Con_L}$ and $\mathsf{Con_R}$ and vice versa. Moreover, since $\mathcal{R}$ contains the congruence rules, applications of the general axioms can be derived in $\mathsf{G'c}\mathcal{R}[\mathsf{Mcon}]$, as is seen by an easy induction on the complexity of the principal formula of the generalised axiom. Thus a sequent is derivable in $\mathsf{Gc}\mathcal{R}[\mathsf{Con}]$ iff it is derivable in $\mathsf{G'c}\mathcal{R}[\mathsf{Mcon}]$. Hence invertibility of the propositional rules in $\mathsf{G'c}\mathcal{R}[\mathsf{Mcon}]$ yields the result for $\mathsf{Gc}\mathcal{R}[\mathsf{Con}]$ as well (but without necessarily preserving the depth of the derivations).

So consider e.g. the rule $\to_{\mathsf{R}}$. We show by induction on the depth of the derivation that $\to_{\mathsf{R}}$ is depth-preserving invertible in $\mathsf{G'c}\mathcal{R}[\mathsf{Mcon}]$, i.e. that if a sequent $\Gamma \Rightarrow \Delta, A \to B$ is derivable in $\mathsf{G'c}\mathcal{R}[\mathsf{Mcon}]$ with a derivation of depth n, then so is the sequent $\Gamma, A \Rightarrow \Delta, B$. If $n = 0$, then the last applied rule was the axiom rule and we have a derivation

$$\frac{}{\Gamma', p \Rightarrow p, \Delta', A \to B}\ \mathcal{A}$$

and thus the formula $A \to B$ cannot have been principal. Thus the sequent $\Gamma', p, A \Rightarrow p, \Delta', B$ is derivable using the axiom rule as well. The case of the last applied rule being $\bot_{\mathsf{L}}$ is similar. If $n = m + 1$, then if the last applied rule was Mcon and the formula $A \to B$ was not part of the contracted sequent, then an application of the induction hypothesis to the premiss followed by an application of Mcon gives the sequent $\Gamma, A \Rightarrow \Delta, B$. If $A \to B$ was part of the contracted sequent we have the situation

$$\dfrac{\begin{array}{c}\mathcal{D}\\ \vdots\\ \Gamma', \Gamma', \Sigma \Rightarrow \Delta', A \to B, \Delta', A \to B, \Pi\end{array}}{\Gamma', \Sigma \Rightarrow \Delta', A \to B, \Pi}\ \mathsf{Mcon}$$

which using the induction hypothesis this is easily turned into

$$\dfrac{\begin{array}{c}\mathcal{D}'\\ \vdots\\ \Gamma', A, \Gamma', A, \Sigma \Rightarrow \Delta', B, \Delta', B, \Pi\end{array}}{\Gamma', A, \Sigma \Rightarrow \Delta', B, \Pi}\ \mathsf{Mcon}\quad .$$

If the last applied rule was the rule $\to_{\mathsf{R}}$ with principal formulae $\Rightarrow A \to B$, then its premiss $\Gamma, A \Rightarrow \Delta, B$ is obviously derivable in depth m. Otherwise the last applied rule was a rule in $\mathcal{R}$ or a propositional rule. Since we can view the propositional rules as rules with restrictions as well and since it is easy to see that the rule set $\mathsf{G'c}$ is Gc-inverting, it suffices to consider the case of a rule $R \in \mathcal{R}$. Since $\mathcal{R}$ is Gc-inverting we know that for every restriction $\mathcal{C}$ from R with $(\Rightarrow A \to B) \restriction_{\mathcal{C}} = \Rightarrow A \to B$ we also have $(A \Rightarrow B) \restriction_{\mathcal{C}} = A \Rightarrow B$. Thus applying the induction hypothesis to the premisses of the application of R, then applying the rule R (possibly together with depth preserving admissibility of weakening) yields a derivation of $\Gamma, A \Rightarrow \Delta, B$ of depth n.

The proofs for the other propositional rules are similar. □

Thus instead of showing mixed-cut closure for the whole set $\mathsf{Gc}\mathcal{R}$ of rules we might be tempted to only show mixed-cut closure and Gc-invertibility for $\mathcal{R}$ and use the latter in the proof of cut elimination whenever the cut formula is principal in a propositional rule and contextual in a rule from $\mathcal{R}$. But interestingly it turns out that if we can prove cut elimination for a sequent system which is sound and complete for a modal logic with non-trivial modalities at all, then invertibility of the propositional rules and mixed-cut closure are equivalent in the sense that we can convert a rule set satisfying one condition into one satisfying the other.

Theorem 2.5.6 (c)**.** *Let $\mathcal{L}$ be a logic based on classical propositional logic such that*

- *the modalities of $\mathcal{L}$ satisfy* congruence, *i.e. for every $n \in \mathbb{N}$ and n-ary modality $\heartsuit$ from Λ we have: whenever $\models_{\mathcal{L}} p_i \leftrightarrow q_i$ for $i \leq n$, then $\models_{\mathcal{L}} \heartsuit(p_1, \dots, p_n) \leftrightarrow \heartsuit(q_1, \dots, q_n)$*
- *the modalities of $\mathcal{L}$ are* non-trivial, *i.e. for every $n \in \mathbb{N}$ and n-ary modality $\heartsuit$ from Λ we have $\not\models_{\mathcal{L}} \heartsuit(p_1, \dots, p_n)$ and $\not\models_{\mathcal{L}} \heartsuit(p_1, \dots, p_n) \to \bot$.*

Then there is a set $\mathcal{R}_1$ of modal rules such that $\mathsf{Gc}\mathcal{R}_1\mathsf{Con}[\mathsf{W}]$ is sound and complete for $\mathcal{L}$ and $\mathsf{Gc}\mathcal{R}_1$ is mixed-cut closed if and only if there is a mixed-cut closed set $\mathcal{R}_2$ of modal rules such that $\mathsf{Gc}\mathcal{R}_2\mathsf{Con}[\mathsf{W}]$ is sound and complete for $\mathcal{L}$ and $\mathcal{R}_2$ is $G\mathsf{c}$-inverting.

Proof. The first observation is that whenever $\mathsf{Gc}\mathcal{R}$ is mixed-cut closed or $\mathcal{R}$ is Gc-inverting, then whenever a rule has a restriction $\langle F_1, F_2 \rangle$ with $p \in F_1$ (resp. $p \in F_2$), then also $p \in F_2$ (resp. $p \in F_1$). This is seen by either applying the criterion of mixed-cut closure to this rule and the propositional rule $\vee_{\mathsf{R}}$ (resp. $\wedge_{\mathsf{L}}$), or the invertibility criterion on this rule and the rule $\rightarrow_{\mathsf{L}}$ (resp. $\rightarrow_{\mathsf{R}}$).

Now we show that for every n-ary modality $\heartsuit$ such a set of rules must contain either a rule R whose principal formulae contain the sequent $\heartsuit(p_1, \ldots, p_2) \Rightarrow \heartsuit(q_1, \ldots, q_n)$, or two rules R_1, R_2 such that

- $\Rightarrow \heartsuit(q_1, \ldots, q_n) \subseteq \mathsf{PF}\,(R_1)$ and $\heartsuit(p_1, \ldots, p_n) \Rightarrow$ satisfies a restriction of R_1
- $\heartsuit(p_1, \ldots, p_n) \Rightarrow \subseteq \mathsf{PF}\,(R_2)$ and $\Rightarrow \heartsuit(q_1, \ldots, q_n)$ satisfies a restriction of R_2.

To see this consider that since the formula $\heartsuit(p_1, \ldots, p_n) \to \heartsuit(p_1 \vee p_1, p_2, \ldots, p_n)$ is $\mathcal{L}$-valid the sequent $\heartsuit(p_1, \ldots, p_n) \Rightarrow \heartsuit(p_1 \vee p_1, p_2, \ldots, p_n)$ must be $\mathsf{Gc}\mathcal{R}\mathsf{Con}$-derivable. Now suppose that in this derivation there is no application of a rule Q with $\Rightarrow \heartsuit(p_1 \vee p_1, p_2, \ldots, p_n) \subseteq \mathsf{PF}\,(Q)$. Then, since no rule decreases the complexity of a formula when passing from the premisses to the conclusion, all the sequents at the leafs of the derivation have the form $\Gamma \Rightarrow \Delta$ with $\mathsf{Supp}\,(\Gamma) \subseteq \{p_1, \ldots, p_n, \heartsuit(p_1, \ldots, p_n)\}$ and $\mathsf{Supp}\,(\Delta) \subseteq \{p_1, \ldots, p_n, \heartsuit(p_1 \vee p_1, p_2, \ldots, p_n)\}$. Thus the formula $\heartsuit(p_1 \vee p_1, p_2, \ldots, p_n)$ cannot have been principal in any of the applications of the axiom rule, and by omitting every occurrence of this formula we obtain a derivation of the sequent $\heartsuit(p_1, \ldots, p_n) \Rightarrow$. But this contradicts non-triviality of the modalities of $\mathcal{L}$. Thus there must be an application of such a rule Q with $\Rightarrow \heartsuit(p_1 \vee p_1, p_2, \ldots, p_n) \subseteq \mathsf{PF}\,(Q)$. If we take the lowermost of such applications, then its conclusion is of the form $\Sigma \Rightarrow \Pi$ with $\mathsf{Supp}\,(\Sigma) \subseteq \{p_1, \ldots, p_n, \heartsuit(p_1, \ldots, p_n)\}$ and $\mathsf{Supp}\,(\Pi) \subseteq \{p_1, \ldots, p_n, \heartsuit(p_1 \vee p_1, p_2, \ldots, p_n)\}$. If $\heartsuit(p_1, \ldots, p_n) \Rightarrow \subseteq \mathsf{PF}\,(Q)$ then we have found the desired rule R. Otherwise one of $\heartsuit(p_1, \ldots, p_n) \Rightarrow$, $p \Rightarrow$ or $\Rightarrow p$ must satisfy a restriction of Q, since otherwise $\Rightarrow \heartsuit(p_1 \vee p_1, p_2, \ldots, p_n)$ would be derivable, in contradiction to non-triviality of the modalities of $\mathcal{L}$. In the first two cases we have found our rule R_1, and in the third case we use the fact shown above that if $\Rightarrow p$ satisfies a restriction, then so does $p \Rightarrow$. In the same way we show the existence of a rule R_2, starting with the formula $\heartsuit(p_1 \wedge p_1, p_2, \ldots, p_n) \to \heartsuit(p_1, \ldots, p_n)$ instead.

Now assume that $\mathsf{Gc}\mathcal{R}_1$ is mixed-cut closed and take a rule $Q \in \mathcal{R}_1$ with restriction $\langle F_1, F_2 \rangle$. If $\heartsuit(A_1, \ldots, A_n) \in F_1$ for any formulae $A_1, \ldots, A_n$ we may apply the condition of mixed-cut closure on this rule and rule R or R_1 from above to get that also $\heartsuit(p_1, \ldots, p_n) \in F_1$, and analogously for $\heartsuit(A_1, \ldots, A_n) \in F_2$ with R or R_2. In the same way using the propositional rules instead of R_1, R_2 we show that if a formula B whose top-level connective is propositional is in F_1 (resp. F_2), then so is p. Thus w.l.o.g. for all restrictions $\langle F_1, F_2 \rangle$ of rules in $\mathcal{R}_1$ we have $F_i \subseteq \{p\} \cup \{\heartsuit(p_1, \ldots, p_n) \mid \heartsuit \in \Lambda \smallsetminus \{\neg, \vee, \wedge, \to, \leftrightarrow\}\}$. But since the propositional connectives are not in Λ it is not hard to see that such a rule set is Gc-inverting.

If on the other hand $\mathcal{R}_2$ is Gc-inverting, then again we get an equivalent rule set with the restriction on the context restrictions of rules in this set, although this case is slightly more involved. The first step is that since $\mathcal{R}_2$ is Gc-inverting, we may permute all applications of propositional rules in a derivation below the applications of modal rules. But then all context formulae in the applications of modal rules are variables or modal formulae, so we may equivalently replace the rule set $\mathcal{R}_2$ with a rule set $\mathcal{R}'_2$ where the context restrictions only contain variables or modalised formulae. Since the set $\mathcal{R}_2$ is mixed-cut closed and contains only modal rules the set $\mathcal{R}'_2$ is mixed-cut closed as well. Now as above using mixed-cut closure and the existence of the rules R resp. R_1 and R_2 from above we get that w.l.o.g. the restrictions contain only variables or modalised variables. Again since the rules in $\mathcal{R}'_2$ are modal rules we obtain that $\mathsf{Gc}\mathcal{R}'_2$ is mixed-cut closed. □

The preceding proof moreover shows that the rules in such a rule set have a very specific form, a very interesting result which is worth stating in its own right, and which we will use extensively when investigating the limits of expressibility of systems given by rules with restrictions in the next chapter.

Corollary 2.5.7 (c)**.** *Let $\mathcal{L}$ be a logic with congruence and non-trivial modalities and let $\mathcal{R}$ be a mixed-cut closed set of modal rules such that $\mathsf{Gc}\mathcal{R}\mathsf{Con}[\mathsf{W}]$ is sound and complete for $\mathcal{L}$ and such that $\mathsf{Gc}\mathcal{R}$ is mixed-cut closed or $\mathcal{R}$ is Gc-inverting. Then w.l.o.g. every restriction of a rule in $\mathcal{R}$ contains only variables or modalised variables.* □

2.6 Dealing with Contraction

In the proof of cut elimination for saturated rule sets we already made use of the property of a rule set being contraction-closed. Intuitively, this allows to permute applications of Con above applications of rules from the rule set whenever the two contracted instances of a formula *both* were principal in the last applied rule. For admissibility of contraction this is not quite enough: we also need to consider the cases where both instances were contextual or where one instance was contextual and one instance was principal in the last applied rule. While the first of these cases can be dealt with in the standard fashion, for the second case we need to slightly modify our sequent calculi. For this we follow Kleene's method for the $\mathsf{G3}$-systems of propositional logic from [Kle52] and not only copy the context, but also the relevant parts of the principal formulae into the premisses. This might be considered a very coarse method, and indeed for classical or intuitionistic propositional logic there are other methods available [TS00]. Unfortunately these methods heavily rely on invertibility of the logical rules, a feature which logical rules for modalities in general do not possess.

Definition 2.6.1. For a rule $R = (\mathcal{P}; \Sigma \Rightarrow \Pi)$ a *modified application*

$$\frac{\{(\Gamma, \Sigma\sigma) \restriction_{F_1}, \Theta\sigma \Rightarrow (\Delta, \Pi\sigma) \restriction_{F_2}, \Upsilon\sigma \mid (\Theta \Rightarrow \Upsilon;\ \langle F_1, F_2 \rangle) \in \mathcal{P}\}}{\Gamma, \Sigma\sigma \Rightarrow \Delta, \Pi\sigma}$$

of R is given by a substitution $\sigma : \mathsf{Var} \to \mathcal{F}$ and a context $\Gamma \Rightarrow \Delta \in \mathcal{S}(\mathcal{F})$. We write $\vdash_{\mathcal{R}^*}$ for derivability using modified applications instead of applications of rules in $\mathcal{R}$.

Example 2.6.2. 1. A modified application of the propositional rule $\to_\mathsf{L}$ has the form

$$\frac{\Gamma, A \to B, B \Rightarrow \Delta \qquad \Gamma, A \to B \Rightarrow \Delta, A}{\Gamma, A \to B \Rightarrow \Delta}$$

2. Modified applications of the rules $\mathsf{4}_n$ have the form

$$\frac{\Box\Sigma, \Box A_1, \ldots, \Box A_n, A_1, \ldots, A_n \Rightarrow B}{\Gamma, \Box\Sigma, \Box A_1, \ldots, \Box A_n \Rightarrow \Box B, \Delta}$$

where none of the formulae in Γ has $\Box$ as its main connective.

3. A modified application of the rule $\{(p_1, p_2 \Rightarrow q; \mathcal{C}_\emptyset), (\Rightarrow q; \mathcal{C}_\mathsf{4})\}/\Box p_1, \heartsuit p_2 \Rightarrow \Box q$ has the form

$$\frac{A_1, A_2 \Rightarrow B \qquad \Gamma\restriction_{\{\Box p\}}, \Box A_1 \Rightarrow B}{\Gamma, \Box A_1, \heartsuit A_2 \Rightarrow \Box B, \Delta}$$

Thus if $\Gamma = \Sigma, \Box\Theta$ where no formula in Σ is of the form $\Box A$ this modified application takes the form

$$\frac{A_1, A_2 \Rightarrow B \qquad \Box\Theta, \Box A_1 \Rightarrow B}{\Sigma, \Box\Theta, \Box A_1, \heartsuit A_2 \Rightarrow \Box B, \Delta}$$

It is not hard to see that the weakening rule is still admissible.

Lemma 2.6.3 (Admissibility of Weakening)**.** *Let $\mathcal{R}$ be a set of rules with restrictions and let $\Gamma \Rightarrow \Delta$ be a sequent. Then we have*

$$\vdash_{\mathcal{R}^*\mathsf{W}[\mathsf{Con}]} \Gamma \Rightarrow \Delta \quad \textit{iff} \quad \vdash_{\mathcal{R}^*[\mathsf{Con}]} \Gamma \Rightarrow \Delta$$

and the depth of the derivations is preserved.

Proof. Analogous to the proof of Lemma 2.3.12 □

Since Weakening is admissible it follows that in the presence of Contraction a sequent is derivable using applications of rules if and only if it is derivable using modified applications of rules.

Proposition 2.6.4. *Let $\mathcal{R}$ be a set of rules with restrictions. Then for every sequent $\Gamma \Rightarrow \Delta$ we have $\vdash_{\mathcal{R}\mathsf{Con}} \Gamma \Rightarrow \Delta$ iff $\vdash_{\mathcal{R}^*\mathsf{Con}} \Gamma \Rightarrow \Delta$.*

Proof. By admissibility of Weakening it is clear that we can derive the premisses of a modified application of a rule from the premisses of the corresponding application. On the other hand given the premisses of a modified application of a rule we simply use a standard application of this rule followed by a number of contractions of the principal formulae. □

The move to modified applications instead of applications of rules allows us to permute contractions between principal and context formulae into the premisses of a rule, thus yielding admissibility of contraction in cut-free sequent systems given by contraction-closed rule sets.

Theorem 2.6.5 (Admissibility of Contraction). *For every contraction-closed set $\mathcal{R}$ of rules and sequent $\Gamma \Rightarrow \Delta$ we have $\vdash_{\mathcal{R}^*\mathsf{Con}} \Gamma \Rightarrow \Delta$ iff $\vdash_{\mathcal{R}^*} \Gamma \Rightarrow \Delta$.*

Proof. The "if" direction is immediate. For the "only if" direction we employ a double induction on the modal nesting depth of the contracted formula and on the depth of the derivation. If the Contraction is applied to an axiom or the conclusion of an application of the left introduction rule for $\bot$ we eliminate it the standard way. So suppose the premiss of the application of the Contraction rule is the conclusion of a rule in $\mathcal{R}$. If the Contraction is between two context formulae or between a context formula and a principal formula, we permute the application of Contraction into the premisses of this rule and eliminate it using the inner induction hypothesis. If Contraction is applied to two principal formulae of a rule R we use contraction closure of the rule set to replace the application of the rule and the Contraction by a number of Contractions and Weakenings on the premisses of that rule and a rule application Q from the rule set. W.l.o.g. all of the newly introduced Contractions are above the newly introduced Weakenings and none of the Contractions is on a context formula of Q. Since the rules add one layer of modalities in the principal formulae, the newly introduced Contractions must be on formulae of lower modal nesting depth and we may eliminate them using the outer induction hypothesis. Finally the applications of Weakening are eliminated using admissibility of Weakening. It is clear from the proof of the latter that this does not introduce any new Contractions. $\square$

Remark 2.6.6. The results of this section show that we can view many sequent calculi which copy the principal formulae into the premisses to ensure admissibility of contraction basically as calculi given by rules with context restrictions if by deleting the copies of the principal formulae from the premisses we obtain rules with context restrictions. This applies e.g. to the sequent calculus $\mathsf{G3s}$ for $\mathsf{S4}$ given in [TS00, p.287].

2.7 Generic Decision Procedures and Complexity

Ultimately, we are interested in deciding for a given formula whether it is a theorem of a particular logic. We call this problem the *validity problem* for a logic.

Validity in $\mathcal{L}$
Input: A formula A
Question: Is $\models_{\mathcal{L}} A$?

From the point of view of the sequent calculi considered here this problem takes the form of deciding whether a given sequent is derivable in a sequent system given by a set of rules with context restrictions. That is, for a given set $\mathcal{R}$ of rules with context restrictions we consider the following decision problem:

Derivability in $\mathcal{R}$
Input: A sequent $\Gamma \Rightarrow \Delta$
Question: Is $\Gamma \Rightarrow \Delta$ derivable in $\mathcal{R}\mathsf{CutConW}$?

Provided we know that the sequent calculus under consideration is saturated, by the generic cut elimination theorem, admissibility of Weakening and the results of the previous section it is clear that we might also equivalently consider the problem of cut-free derivability in the modified rule set $\mathcal{R}^*$:

Cut-free Derivability in $\mathcal{R}^*$
Input: A sequent $\Gamma \Rightarrow \Delta$
Question: Is $\Gamma \Rightarrow \Delta$ derivable in $\mathcal{R}^*$?

Before taking a closer look at generic decidability and complexity results for these problems we briefly recapitulate different representations and measures of size of formulae. The standard representation of formulae is as a string of symbols.

Definition 2.7.1. Let Λ be a finite set of connectives with associated arities and let F be a formula with connectives in Λ. The *formula size of* F is the number $\|F\|$ of symbols in F.

We will also consider the more succinct circuit (or DAG-) representation of formulae. This representation allows us to identify different occurrences of the same subformula.

Definition 2.7.2. Let Λ be a finite set of connectives with associated arities. A Λ-*circuit* is a directed acyclic graph whose nodes are labelled in the following way:

1. Nodes without any predecessors are called *input nodes* and are labelled with a propositional variable or a 0-ary connective from Λ.

2. Nodes with $k > 0$ predecessors are labelled with a k-ary connective from Λ.

We assume that the predecessors of every node are ordered. Nodes without any successors are called *output nodes*. The *size* of a Λ-circuit C is the number $s(C)$ of nodes in C. A Λ-circuit with one output node corresponds in the obvious way to a formula with connectives in Λ. For such a formula F the *circuit size of* F is the minimal size of a Λ-circuit representing F. The circuit size of the formula F is denoted by $\|F\|_c$. For a sequent $\Gamma \Rightarrow \Delta$ the *circuit size of* $\Gamma \Rightarrow \Delta$ is defined as $\|\Gamma \Rightarrow \Delta\|_c := \sum_{F \in \mathsf{Supp}(\Gamma,\Delta)} \|F\|_c$.

Thus the circuit size of a formula is the number of different subformulae occurring in it. Using e.g. the formulae from the family of formulae $\{A_n \mid n \in \mathbb{N}\}$ defined by $A_0 := p$ and $A_{n+1} := A_n \wedge A_n$ for $n \in \mathbb{N}$ it can be seen that the circuit presentation of formulae is potentially exponentially more succinct than the standard representation. The general idea for our generic decision procedure now is the following: Given a saturated set of rules with restrictions we know by Theorem 2.4.16 that the cut rule is admissible and thus safely can be dropped. Then we modify the sequent system as described in the previous section, yielding admissibility of contraction. Furthermore, by admissibility of Weakening the Weakening rule can be dropped as well. Now we would like to either use the subformula property of the resulting sequent system to enumerate all sequents possibly occurring in the derivation of a given sequent, or we would like to try backwards proof search to find a derivation for it. In general our rule sets will comprise infinitely many rules, though, so we will need to impose another condition on the rule sets which ensures that the necessary rules can be computed fast enough.

Definition 2.7.3. We write A^* for the set of finite strings of symbols from a set A. For finite sets A, B a relation $R \subseteq A^* \times B^*$ is called PSPACE-*tractable* if given a tuple $(a, b) \in A^* \times B^*$ it is decidable in space polynomial in the length of a whether $(a, b) \in R$. A set $\mathcal{R}$ of rules is *tractable* if there exists an encoding of the applications of rules in $\mathcal{R}$ such that both the relation holding between sequents and codes of applications of rules with this sequent as conclusion and the relation holding between codes of applications of rules and their premisses are PSPACE-tractable.

Example 2.7.4. The rule sets $\mathsf{G[cim]}$ as well as $\mathcal{R}_{\mathsf{K}}, \mathcal{R}_{\mathsf{KT}}, \mathcal{R}_{\mathsf{K4}}$ and $\mathcal{R}_{\mathsf{S4}}$ are tractable.

Indeed, if a saturated rule set is tractable, then we can use a generic algorithm to decide whether a sequent is derivable using modified applications.

Theorem 2.7.5. *Let $\mathcal{R}$ be a saturated and tractable set of rules with restrictions. Then the cut-free derivability problem for $\mathcal{R}^*$ is in* EXPTIME. *More precisely, it is decidable in time exponential in the circuit size of the input.*

Proof. The idea is to work on fully contracted sequents and make use of the subformula property of the system to compute all the derivable sequents built from subformulae of the original sequent. Since all formulae occurring in the premisses of modified applications of rules in $\mathcal{R}$ are subformulae of the formulae occurring in the conclusion of this application, it is clear that if a sequent is derivable, then it has a derivation in which only such sequents occur (although not necessarily fully contracted).

Recall that for a multiset Γ of formulae we write $\mathsf{Supp}\,(\Gamma)$ for the *support* of Γ, that is the multiset of formulae in Γ disregarding their multiplicities. Furthermore, for a sequent $\Gamma \Rightarrow \Delta$ let $\mathsf{Sf}(\Gamma \Rightarrow \Delta)$ denote the set of subformulae of $\Gamma \Rightarrow \Delta$, where as usual we identify different occurrences of the same formula. Then obviously we have $|\mathsf{Sf}(\Gamma \Rightarrow \Delta)| \leq \|\Gamma \Rightarrow \Delta\|_c$. Since the rule set is contraction closed, by Theorem 2.6.5 the contraction rule is admissible, and by Lemma 2.6.3 the weakening rule is admissible as well. Thus it is clear that a sequent $\Gamma \Rightarrow \Delta$ is derivable in $\mathcal{R}^*$ if and only if the sequent $\mathsf{Supp}\,(\Gamma) \Rightarrow \mathsf{Supp}\,(\Delta)$ is derivable in $\mathcal{R}^*$. Let

$\mathfrak{S}(\Gamma \Rightarrow \Delta)$ denote the set of sequents $\mathsf{Supp}\,(\Sigma) \Rightarrow \mathsf{Supp}\,(\Pi)$ with $\Sigma \Rightarrow \Pi \in \mathcal{S}(\mathsf{Sf}(\Gamma \Rightarrow \Delta))$. The procedure given as Algorithm 1 checks derivability in $\mathcal{R}^*$ on input $\Gamma \Rightarrow \Delta$ by first iteratively constructing all relevant derivable sequents and then checking whether $\Gamma \Rightarrow \Delta$ is in this set.

Algorithm 1: Decision procedure for derivability in $\mathcal{R}^*$

Input: a sequent $\Gamma \Rightarrow \Delta$
set $\mathcal{D}, \mathcal{D}' := \emptyset$;
repeat
  set $\mathcal{D} := \mathcal{D}'$;
  foreach $\Sigma \Rightarrow \Pi \in \mathfrak{S}(\Gamma \Rightarrow \Delta)$ **do**
    if *exists application of* $R \in \mathcal{R}^*$ *with conclusion* $\Sigma \Rightarrow \Pi$ *s.t. for all premisses* $\Theta \Rightarrow \Xi$ *of this application of* R*:* $\mathsf{Supp}\,(\Theta) \Rightarrow \mathsf{Supp}\,(\Xi) \in \mathcal{D}$ **then**
      add $\Sigma \Rightarrow \Pi$ to $\mathcal{D}'$
until $\mathcal{D} = \mathcal{D}'$;
$\Gamma \Rightarrow \Delta$ is derivable iff $\Gamma \Rightarrow \Delta \in \mathcal{D}$

Since $\mathcal{D} \subseteq \mathcal{D}' \subseteq \mathfrak{S}(\Gamma \Rightarrow \Delta)$ and since the number of sequents in $\mathfrak{S}(\Gamma \Rightarrow \Delta)$ is only exponential in $|\mathsf{Sf}(\Gamma \Rightarrow \Delta)| =: s$, the **repeat**-loop in the procedure is executed at most exponentially often (in s). Furthermore, in each execution of the loop the procedure checks at most exponentially many sequents, and since the rule set is tractable, checking each sequent can be done in time exponential in s. Thus the overall runtime of the procedure is exponential in the number of subformulae of the input sequent. Since $|\mathsf{Sf}(\Gamma \Rightarrow \Delta)| \leq \|\Gamma \Rightarrow \Delta\|_c$ this yields the result. □

Corollary 2.7.6. *Let $\mathcal{R}$ be a saturated and tractable set of rules with restrictions. Then the derivability problem for $\mathcal{R}$ is in* EXPTIME.

Proof. We have for every sequent $\Gamma \Rightarrow \Delta$:

$$\vdash_{\mathcal{R}\mathsf{CutConW}} \Gamma \Rightarrow \Delta \overset{\text{Lem. 2.3.12}}{\Longleftrightarrow} \vdash_{\mathcal{R}\mathsf{CutCon}} \Gamma \Rightarrow \Delta \overset{\text{Thm. 2.4.16}}{\Longleftrightarrow} \vdash_{\mathcal{R}\mathsf{Con}} \Gamma \Rightarrow \Delta$$
$$\overset{\text{Prop. 2.6.4}}{\Longleftrightarrow} \vdash_{\mathcal{R}^*\mathsf{Con}} \Gamma \Rightarrow \Delta \overset{\text{Thm. 2.6.5}}{\Longleftrightarrow} \vdash_{\mathcal{R}^*} \Gamma \Rightarrow \Delta\ .$$

Together with Theorem 2.7.5 this yields the result. □

Corollary 2.7.7. *Since the rule sets $\mathcal{R}_{\mathsf{K4}}$ and $\mathcal{R}_{\mathsf{S4}}$ are as we have seen saturated and tractable, their derivability problems are in* EXPTIME.

Of course it is well-known that the logics $\mathsf{K4}$ and $\mathsf{S4}$ are decidable in PSPACE [Lad77], and so the complexity bounds obtained in the previous corollary are not optimal. This raises the question whether we can do better in general. Since the rules for modal logic K are rules with restrictions, and since the problem of deciding whether a formula is valid in K is known to be PSPACE-complete [Lad77], it is clear that the complexity of a generic decision procedure will not be below PSPACE. It is not known whether in general the EXPTIME complexity is optimal.

On the other hand, if we restrict the rule format to shallow rules it *is* possible to bring the generic bound down to the optimal complexity.

Theorem 2.7.8. *Let $\mathcal{R}$ be a saturated and tractable set of shallow rules. Then the cut-free derivability problem for $\mathcal{R}^*$ is in* PSPACE. *More precisely, it can be solved in space polynomial in the circuit size of the input.*

Proof. In a first step we note that since Weakening and Contraction are admissible we might equivalently work with sequents based on *sets* instead of multisets and restrict modified applications of rules in such a way that all the principal formulae are distinct. Then it can be easily seen that for every set-sequent derivable in this system there is a derivation in which on every branch every set-sequent occurs at most once. Furthermore, since every formula occurring in the premisses of a rule is a subformula of a formula occurring in its conclusion, only set-sequents built from subformulae of a set-sequent can occur in its derivation. Now backwards proof search for this system is implemented on an alternating Turing machine as follows: the machine existentially guesses the last applied rule, then universally guesses its premisses (both of which only take polynomial space since the rule set is tractable) and checks that the premisses have not been encountered before and are derivable. Due to the format of the rules when passing from conclusion to premisses either the set-sequent is extended by at least one formula or the maximal complexity of formulae is diminished by at least one. Since both the number of formulae in the set-sequents and their maximal complexity are bounded by the number of subformulae of the end-sequent and thus the circuit-size of the input sequent, the branches in the computation tree have length polynomial in the latter value. Thus the procedure runs in alternating polynomial time, which is equivalent to polynomial space [CKS81]. □

Then as above we have:

Corollary 2.7.9. *Let $\mathcal{R}$ be a saturated and tractable set of shallow rules. Then the derivability problem for $\mathcal{R}$ is in* PSPACE. □

In particular, this yields uniform decision procedures of optimal complexity for a wide variety of standard modal logics and intuitionistic propositional logic.

Example 2.7.10.

1. As we have seen above the rule sets $\mathsf{G[cim]}$ are saturated and tractable. Thus the derivability problems for these rule sets are in PSPACE.
2. Similarly, the rule sets $\mathcal{R}_{\mathsf{K}}$ and $\mathcal{R}_{\mathsf{KT}}$ are saturated and tractable. Thus the derivability problems for $\mathsf{Gc}\mathcal{R}_{\mathsf{K}}$ and $\mathsf{Gc}\mathcal{R}_{\mathsf{KT}}$ are in PSPACE.

Example 2.7.11. We can use this generic decision procedure to solve the derivability problem for a number of extensions of *constructive modal logic* CMK (see e.g. [Wij90, BdPR01]). The rules and rule sets for the logics CMK and CMKT are given in Table 2.6. While these are not

$$\begin{aligned}
\mathsf{K}\Box_n &:= \{(p_1, \dots, p_n \Rightarrow q; \mathcal{C}_\emptyset)\} / \Box p_1, \dots, \Box p_n \Rightarrow \Box q \\
\mathsf{K}\Diamond_n &:= \{(p_1, \dots, p_n, q \Rightarrow r; \mathcal{C}_\emptyset)\} / \Box p_1, \dots, \Box p_n, \Diamond q \Rightarrow \Diamond r \\
R_{\mathsf{T}\Box} &:= \{(p \Rightarrow ; \mathcal{C}_{\mathsf{id}}) / \Box p \Rightarrow \\
R_{\mathsf{T}\Diamond} &:= \{(\Rightarrow p; \mathcal{C}_{\mathsf{id}}) / \ \Rightarrow \Diamond p
\end{aligned}$$

$$\begin{aligned}
\mathcal{R}_{\mathsf{CMK}} &:= \{\mathsf{K}\Box_n \mid n \geq 0\} \cup \{\mathsf{K}\Diamond_n \mid n \geq 0\} \\
\mathcal{R}_{\mathsf{CMKT}} &:= \mathcal{R}_{\mathsf{CMK}} \cup \{R_{\mathsf{T}\Box}, R_{\mathsf{T}\Diamond}\}
\end{aligned}$$

Table 2.6: The rule sets for some constructive modal logics

yet principal-cut closed (albeit principal-cut deriving, see Remark 2.4.18), by Lemma 2.4.5 we may simply add the missing cuts between rules $\mathsf{K}\Box_n$ and $R_{\mathsf{T}\Box}$ resp. $R_{\mathsf{T}\Diamond}$ and $\mathsf{K}\Diamond_n$ to the rule set. This gives the additional rules

$$\begin{aligned}
\mathsf{T}\Box_n &:= \{(p_1, \dots, p_n \Rightarrow ; \mathcal{C}_{\mathsf{id}})\} / \Box p_1, \dots, \Box p_n \Rightarrow \\
\mathsf{T}\Diamond_n &:= \{(p_1, \dots, p_n \Rightarrow q; \mathcal{C}_{\mathsf{id}})\} / \Box p_1, \dots, \Box p_n \Rightarrow \Diamond q
\end{aligned}$$

and the rule set $\mathcal{R}'_{\mathsf{CMKT}} := \mathcal{R}_{\mathsf{CMK}} \cup \{\mathsf{T}\Box_n \mid n \geq 1\} \cup \{\mathsf{T}\Diamond_n \mid n \geq 0\}$. It is clear that a sequent is derivable in $\mathsf{Gi}\mathcal{R}_{\mathsf{CMKT}}[\mathsf{CutCon}]$ iff it is derivable in $\mathsf{Gi}\mathcal{R}'_{\mathsf{CMKT}}[\mathsf{CutCon}]$. Furthermore, it can be seen that the rule sets $\mathcal{R}_{\mathsf{CMK}}$ and $\mathcal{R}'_{\mathsf{CMKT}}$ are saturated and tractable, and thus by Corollary 2.7.9 derivability in each of the systems $\mathsf{Gi}\mathcal{R}_{\mathsf{CMK}}\mathsf{CutConW}$ and $\mathsf{Gi}\mathcal{R}_{\mathsf{CMKT}}\mathsf{CutConW}$ is decidable in PSPACE.

We will make extensive use of the generic decidability results when we construct complexity optimal sequent calculi for a number of conditional logics in Chapters 4 and 5.

2.8 Notes

The Rule Format. The format of rules with context restrictions is an extension of the format of one-step rules considered e.g. in [PS08, PS10]. The additional feature of rules with context restrictions, the ability to copy the whole or part of the context into the premisses, allows to capture not only standard propositional rules but also rules for modal logics axiomatised by axioms of modal rank greater than one, as we will see in the next chapter. The intermediate format of shallow rules was introduced in [LP11] and the present format has been published in [LP13a].

The notion of a context restriction is inspired by the more general notion of a *context relation* introduced in [AL11]. In contrast to context restrictions, which only restrict the context but do not change it, context relations are arbitrary (finite) binary relations between signed formulae, interpreted by allowing the context formulae to change from premiss to conclusion according to the relation. This not only allows adding modalities to context formulae in the conclusion, thus e.g. interpreting the boxed formulae on the left hand side of the conclusion of the K_n-rules as context formulae, or stripping the context formulae of modalities, thus capturing e.g. the rule for symmetry given in Example 2.3.7 (5), but also moving context formulae from one side of the sequent to the other. While obviously context restrictions can be seen as severely restricted context relations, the additional expressive power of the latter would be counterbalanced by the need for increasingly complex extensions of the syntactic criteria for cut elimination given in Section 2.4. Moreover, the rule format considered in [AL11], that of *basic rules* allows arbitrary sequents as the principal formulae. While this has the advantage of capturing the structural rules as well, it has the disadvantage that it also allows trivially cut-free sequent calculi where for each theorem of a logic there is a rule whose principal formulae simply state this theorem on the right hand side. While this is not a problem for the construction of semantics from a rule set as introduced in [AL11], for our syntactical investigations into cut elimination and also for the questions considered in the next chapter this is clearly not desirable. Of course again every rule with context restrictions can be seen to be a rule with context relations, which makes all the rule sets considered here amenable to the semantic methods based on basic rules. It would be very interesting to see whether the semantic criteria for cut admissibility given in [AL11] and our syntactic criteria for cut elimination are related in some way. Also of course we would like to extend the notion of context restriction towards the notion of context relation in order to be able to capture more logics.

The main difference to the other rule formats studied in the literature is that our rule format allows for several principal formulae, while usually only one principal formula is introduced. However, if we want to stay in the classical sequent framework and not introduce new structural connectives, and if we take the context to be unchanging, then introducing several principal formulae is essential for capturing modal logics, since otherwise already modal logic K would be problematic. One line of research in this direction is based on the notion of a *canonical rule*, which in its basic form [AL01] in our terminology essentially is a shallow rule with restrictions $\mathcal{C}_{\mathsf{id}}$ whose principal formulae contain exactly one formula (the rules in the original form are context-independent while our rules are context-sharing in the sense of [TS00]). Extensions of this format to quantifiers and labelled calculi have subsequently been considered in [ZA06, AZ08, ZA12]. The (asymmetric) *simple* calculi investigated in [CT06b] and the (symmetric) *standard* calculi from [CT06a] also only allow one principal formula per rule, but allow for more differentiations in the context. This is necessary since, in contrast to the calculi considered here, the calculi investigated in these works do not need to contain all the structural rules. The rules for *cut suitable calculi* considered in [Ras07] are slightly more general in that they also allow restrictions of (parts of) the context to formulae with a specific main connective and cardinality restrictions, thus rendering the format suitable for capturing e.g. intuitionistic S4. But even here we only have one principal formula. The rule formats in [CT06a] and [Ras07] also allow for the treatment of quantifiers.

Criteria for Cut Elimination. Our notion of a principal-cut closed rule set is based on the notion of a *resolution closed* rule set in [SP09] and the notion of a *cut absorbing* rule set in [PS10]. The technique of variable elimination given in Definition 2.4.1 and the method of proving that cuts between rules are sound (Lemma 2.4.5) are similar to the method of *cutting* applied to quasiequations in the construction of completed structural rules given in [CGT08, CST09, CGT12]. The operations of cut between rules and (right- resp. left-) contraction of a rule can be seen as admissible rules in a calculus for admissible rules in the spirit of [IM09].

Our proof for cut elimination (Theorem 2.4.16) is based on the proof in [SP09, PS08, PS10] for sequent systems based on one-step rules. A first extension to shallow rules using the multicut rule was stated in [LP11], where unfortunately the necessary condition of contraction-closure was omitted. The generalisation to right- or left-contraction closed rule sets is based on the proof in [vP01]. Where the latter proof relied on invertibility of the propositional rules we make use of context-cut closure and mixed-cut closure of the rule set.

In the literature there exist a number of results stating sufficient and in some cases also necessary conditions for different variants of cut elimination or cut admissibility. In contrast to semantically motivated criteria which are based on providing a generic semantics for logics given by sequent calculi of a specific format [AL09, AL11, BLZ12] we are mainly interested in syntactical criteria. The prime example for this kind of criteria are Belnap's conditions $C1$ to $C8$ ensuring cut elimination for display logic [Bel82]. Condition $C8$ roughly corresponds to our condition of principal-cut closure. Similar criteria for consecution calculi are given in [Res00]. The calculi considered in these works introduce additional structural connectives. Considering the standard sequent format without any additional structure the criterion of principal-cut closure for calculi allowing more than one principal formula was formulated in [PS08, PS10]. But since in contrast to the one-step rules considered there our rules might copy the whole or parts of the context, we need the criteria of context-cut closure and mixed-cut closure as well. The criteria in [PS09, PS11] also demand that the logical rules absorb the structural rules and inversions of the propositional rules. This way the structural rules are admissible and the propositional rules invertible, a fact which the proof of cut elimination presented there relies on. Since the rule format in [Ras07] allows for context-independent rules and since the calculi considered there might not contain all the structural rules, the criteria for cut elimination in this work are considerably more fine grained and involved. Condition $C9$ roughly corresponds to our condition of principal-cut closure (for rules with one principal formula). The conditions of mixed-cut closure and context-cut closure are split into several cases. This work also includes an analysis of the complexity of cut elimination. A sufficient and necessary condition for cut admissibility for canonical calculi was considered in [AL01, ZA06] under the name of *coherence*. This demands that whenever cut can be applied to the conclusions of two canonical rules, then their premisses must be propositionally inconsistent, and in this setting amounts to principal-cut closure of the rule set.

Stronger notions of cut elimination such as *reductive* cut elimination [CT06b], *modular* cut elimination [CT06a] or *strong* cut elimination [AZ08] are based on the idea that whenever a sequent is derivable from a set $\mathcal{S}$ of sequents, in the case of [CT06a] with particular properties, then it must be cut-free derivable from the $\mathcal{S}$ [CT06a] or derivable using cuts only on formulae from $\mathcal{S}$. Necessary and sufficient conditions for these notions of cut elimination in the frameworks of canonical rules, simple rules or standard calculi have been established in the above cited works and are all similar to the coherence condition of [AL01]. In [CT06b, CT06a] the corresponding condition, *reducibility*, also takes into account different format of the context in the rules considered there. Furthermore, this format also dictates the condition of *(weak) substitutivity*, which ensures that cuts with context formulae can be permuted into the premisses of a rule.

Contraction. The method of copying the (relevant) principal formulae into the conclusions of a rule to ensure admissibility of Contraction seems to have been introduced by Kleene for the construction of the $\mathsf{G3}$-systems for propositional logic [Kle52]. As a general method in the construction of cut-free sequent calculi it has been used e.g. in [NvP01, Neg05], where also the idea of a contraction-closed rule set seems to have been explicitly formulated for the first time as the *closure condition*. Our notion of contraction-closed rule sets is based on [SP09, PS10].

Decidability. Our condition on tractability of the rule sets is an extension of the notion of a *pspace-tractable* rule set in [SP09], where a generic PSPACE-decision procedure for modal logics given by one-step rules based on backwards proof search was given. This generic procedure also has been implemented [CMPS09]. Theorem 2.7.8 on decidability for shallow rule sets in PSPACE was originally published in [LP11] with a slightly different proof using histories in the spirit of [HSZ96] to prevent multiple applications of the same rule with restriction $\mathcal{C}_{\mathsf{id}}$ in the backwards proof search algorithm. We will see in the next chapter that logics axiomatised by non-iterative axioms are closely connected to shallow rules. A generic semantic method along with semantic criteria ensuring decidability in PSPACE for such non-iterative modal logics has been given in [SP08]. Decidability for every logic axiomatised by finitely many non-iterative axioms has also been shown in [Lew74] using semantical methods. While not explicitly stated this proof seems to suggest a 3EXPTIME upper bound for such logics. For normal modal logics this result was strengthened in [tC05] to a NEXPTIME upper bound for logics defined by *shallow* formulae, where in such a formula every occurrence of a propositional variable is in the scope of at most one modality. It is not clear whether either of these bounds is tight. In particular, the author is not aware of any finitely axiomatised non-iterative logics with complexity above PSPACE. Since moreover standard proofs for EXPTIME-hardness e.g. via reductions from some form of tiling problem [BdRV01] or acceptance problem for Turing machines seem to rely on nested modalities, this raises the question whether nested axioms are essential for EXPTIME-hardness.

Problem 2.8.1. *Is there an* EXPTIME-*hard finitely axiomatised non-iterative logic?*

3 Axioms versus Rules

In the preceding chapter we have defined a fairly general format of sequent rules and seen criteria on sets of rules of this format which are sufficient to ensure that the sequent calculus given by the set of rules allows for a syntactic proof of cut elimination or can be used in a generic decision procedure. Now we take a closer look at Question 1.1.4 from the Introduction, that is we would like to know for which kinds of modal logics there can be a sound and complete sequent system consisting of rules with context restrictions. For this purpose as mentioned we will take the modal logics to be given by a set of axioms for a *Hilbert system*. Thus we would like to know which axioms can be captured by our rule format, i.e. have corresponding rules with restrictions, and moreover which axioms *can not* be captured.

For this purpose in Section 3.2 we will first syntactically characterise the class of *translatable axioms* and give a purely syntactic and automatic translation of axioms of this class into so-called *proto rules*, that is rules with restrictions where the *number* of context formulae is fixed. This is then used to show that *ω-sets* for translatable axioms, i.e., sets of axioms which are generated in a particular way by a single translatable axiom, are equivalent to a single rule with restrictions. Furthermore, under certain additional restrictions, namely *normality of the context formulae* this generated set of axioms is equivalent to the single generating axiom, and thus the latter can be translated into a rule with restrictions instead of a proto rule. Since the translation is automatic we can take it as a convenient starting point if we try to construct a cut-free sequent system for a modal logic given by a set of axioms in the spirit of Question 1.1.3.

Of course the translation from axioms to rules gives only sufficient conditions for an axiom to be equivalent to a rule with restrictions. Naturally, we would like to know whether these criteria are necessary as well. To address this point we will consider in Section 3.3 a translation from rules with restrictions back to axioms (or sets of axioms) of our format. This will establish necessity of the criteria and thus the following strict correspondence: an axiom for a Hilbert system can be captured by rules with restrictions if and only if it is axiomatically equivalent to a set of ω-sets for translatable axioms. Restrictions of the rule format to one-step or shallow rules yield suitably restricted classes of axioms. As a graphical guide to the translations, the main results are given schematically in Figure 3.1, where arrows indicate translations and are annotated with the corresponding theorems. Dashed arrows indicate that the translations presuppose normal context formulae resp. restrictions. For both directions the main work lies in the translation between translatable clauses and proto rules.

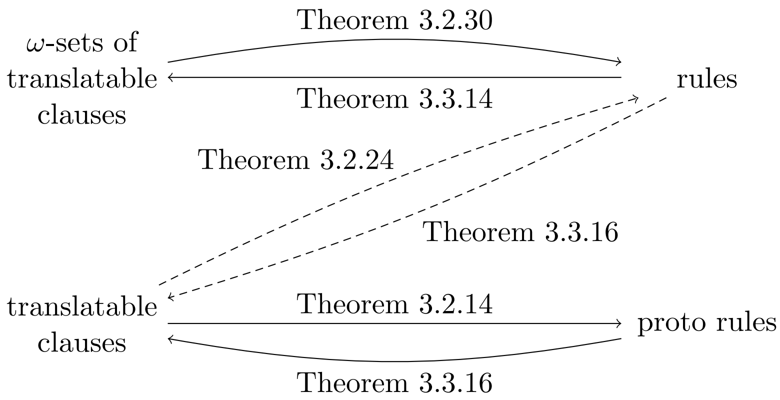


Figure 3.1: The main translations schematically. Dashed lines presuppose normality of the context formulae / restrictions.

In Section 3.4 finally we will make use of this correspondence to show limitative results concerning the expressivity of the different rule formats under consideration. Apart from formally establishing that the three formats of rules have different expressive power this will also show that certain standard modal logics cannot be captured by one-step or shallow rules. It should be noted, however, that the translations make use of the cut rule, and therefore might be a bit blunt if we are interested in limitative results for *cut-free* sequent systems. In particular, for logics such as $\mathsf{S5}$, for which there *are* complete (with Cut) calculi given by rules with restrictions, we will not be able to use the translation to show that they cannot be captured by a cut-free system given by rules with restrictions. Nonetheless, we will see that if we consider sequent calculi given by a *mixed-cut closed* set of rules with restrictions, then we *can* use the translation to show limitative results regarding $\mathsf{S5}$ and a number of other standard modal logics. The assumption of mixed-cut closure seems to be reasonably mild. In particular, if we are interested in rule sets for which cut elimination can be shown in the standard way, mixed-cut closure, i.e. the ability to permute cuts where the cut formula is contextual in one rule into the premisses of this rule, seems to be fundamental. Of course the limitative results themselves are independent of a specific technique for proving cut elimination and thus also preclude e.g. semantic proofs of cut admissibility in such systems.

3.1 Hilbert Axioms and Sequent Rules

Let us first recall the definition and some basic properties of Hilbert systems.

Definition 3.1.1. For a set $\mathcal{A} \subseteq \mathcal{F}(\Lambda)$ of formulae we take the (classical resp. intuitionistic resp. minimal) *Hilbert system* $\mathcal{H}[\mathsf{cim}]\mathcal{A}$ to include the formulae in $\mathcal{A}$ and axioms for the propositional base logic and to be closed under the rule of uniform substitution US, modus ponens MP and the congruence rules $\mathsf{Cong}_\heartsuit$ for all modalities $\heartsuit \in \Lambda$ as given in Table 3.1, see

$$\frac{\vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} A_1 \leftrightarrow B_1 \quad \dots \quad \vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} A_n \leftrightarrow B_n}{\vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} \heartsuit(A_1,\dots,A_n) \leftrightarrow \heartsuit(B_1,\dots,B_n)}\ \mathsf{Cong}_\heartsuit \qquad \frac{\vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} A \qquad \vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} A \to B}{\vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} B}\ \mathsf{MP}$$

$$\frac{\vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} A_1 \to B_1 \quad \dots \quad \vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} A_n \to B_n}{\vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} \heartsuit(A_1,\dots,A_n) \to \heartsuit(B_1,\dots,B_n)}\ \mathsf{Mon}_\heartsuit \qquad \frac{\vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} A}{\vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} A\sigma}\ \mathsf{US}$$

Table 3.1: The derivation rules for Hilbert systems

also [HC96, p.222]. For axiomatisations of the propositional base logics see e.g. [TS00, p.51]. We write $\vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} A$ if the formula A is in $\mathcal{H}[\mathsf{cim}]\mathcal{A}$ and say that A is *derivable* in the Hilbert system $\mathcal{H}[\mathsf{cim}]\mathcal{A}$.

It is clear that Hilbert systems are logics in the sense of Definition 2.1.5. The main difference is that Hilbert systems also are closed under the congruence rules and explicitly mention an axiomatisation. Instead of working directly with Hilbert systems we will work with sequent systems for propositional logic with the rules of congruence and additional "axioms" given by a sequent consisting of only one formula:

Definition 3.1.2. For a set $\mathcal{R}$ of sequent rules and a set $\mathcal{A} \subseteq \mathcal{F}(\Lambda)$ of formulae we write $\mathcal{R} + \mathcal{A}$ for the sequent system consisting of the rules in $\mathcal{R}$ together with the *ground sequents* $\Rightarrow A$ for every $A \in \mathcal{A}$. Ground sequents are treated as zero-premiss rules. In particular all their substitution instances are derivable.

Proposition 3.1.3 (cim)**.** *Let $\mathcal{A} \subseteq \mathcal{F}(\Lambda)$ be a set of axioms. Then for every sequent $\Gamma \Rightarrow \Delta$ we have $\vdash_{\mathsf{G[cim]CongWCutCon}+\mathcal{A}} \Gamma \Rightarrow \Delta$ iff $\vdash_{\mathcal{H}[\mathsf{cim}]\mathcal{A}} \bigwedge \Gamma \to \bigvee \Delta$.*

Proof. The proof is a simple extension of the standard proof for equivalence of sequent systems and Hilbert system for propositional logic [Gen35, Joh37, TS00, NvP01]. For the congruence rules in the Hilbert system we have counterparts in the sequent system, and since an axiom $A \in \mathcal{A}$ for the Hilbert system corresponds directly to a sequent $\Rightarrow A$ the additional axioms do not create any problems. □

In view of the previous Proposition in the following we will concentrate on sequent systems with ground sequents consisting of a single formula on the right hand side instead of axioms for a Hilbert system. The following definition allows us to compare rules and axioms. In order to allow for a background theory we consider the notion of equivalence *over* a set of rules, where the latter of course might be empty.

Definition 3.1.4. Let $\mathcal{R}$ be a set of rules with context restrictions and assume we have fixed a propositional base logic. A set $\mathcal{R}_{\mathcal{A}}$ of rules is *equivalent over* $\mathcal{R}$ to a set $\mathcal{A}$ of axioms if every rule in $\mathcal{R}_{\mathcal{A}}$ is a derivable rule in $\mathsf{G[cim]}\mathcal{R}\mathsf{CutConW} + \mathcal{A}$ and every axiom in $\mathcal{A}$ is derivable in $\mathsf{G[cim]}\mathcal{R}\mathsf{CutConW}\mathcal{R}_{\mathcal{A}}$. In case $\mathcal{R}_{\mathcal{A}} = \{R\}$ and $\mathcal{A} = \{A\}$ we also say that the rule

R is equivalent to the axiom A. Similarly, two sets $\mathcal{R}_1, \mathcal{R}_2$ of rules are *equivalent over* $\mathcal{R}$, if every rule in $\mathcal{R}_1$ is a derivable rule in $\mathsf{G[cim]}\mathcal{R}\mathsf{CutConW}\mathcal{R}_2$ and vice versa. Finally, two sets $\mathcal{A}_1, \mathcal{A}_2$ of axioms are *(axiomatically) equivalent over* $\mathcal{R}$ if every axiom in $\mathcal{A}_1$ is derivable in $\mathsf{G[cim]}\mathcal{R}\mathsf{CutConW} + \mathcal{A}_2$ and vice versa.

This allows us to phrase the main question for this chapter more formally:

Question 3.1.5. Sets of axioms of which syntactical form are equivalent to sets of rules with context restrictions?

In order to have some tools at our disposal we make the following assumption.

Assumption 3.1.6. For the rest of this chapter we assume that the logics are extensions of intuitionistic or classical propositional logic.

3.2 From Axioms to Rules

Before stating the formal definition of the class of translatable axioms let us have a brief look at the general idea for the translation from axioms to rules with restrictions. The main idea here is to consider axioms which are particular substitution instances of axioms without iterated modalities. In a first step then the underlying non-iterative axiom is translated into a rule. Then both in the premisses and the conclusion the variables of the non-iterative axiom are substituted according to the original axiom. Under certain circumstances these formulae behave as context formulae for the resulting sequent rule. Let us have a look at an example.

Example 3.2.1. We take $\heartsuit$ and $\spadesuit$ to be unary modalities for a modal logic based on intuitionistic propositional logic and we stipulate that the rule set $\mathcal{R}$ describing the background theory includes rules ensuring normality of $\spadesuit$ and monotony and "seriality" of $\heartsuit$:

$$\mathcal{R} = \{\ \{(p, q \Rightarrow r; \mathcal{C}_\emptyset)\}/\spadesuit p, \spadesuit q \Rightarrow \spadesuit r,\ \ \{(\Rightarrow p; \mathcal{C}_\emptyset)\}/ \Rightarrow \spadesuit p,\ \ \{(p \Rightarrow\ ; \mathcal{C}_\emptyset)\}/\heartsuit p \Rightarrow\ ,\ \ \mathsf{Mon}_\heartsuit\ \}\ .$$

In the classical notation these rules are given by:

$$\frac{r, q \Rightarrow r}{\Gamma, \spadesuit p, \spadesuit q \Rightarrow \spadesuit r, \Delta} \qquad \frac{\Rightarrow p}{\Gamma \Rightarrow \spadesuit p, \Delta} \qquad \frac{p \Rightarrow}{\Gamma, \heartsuit p \Rightarrow \Delta} \qquad \frac{p \Rightarrow q}{\Gamma, \heartsuit p \Rightarrow \heartsuit q, \Delta}\ \mathsf{Mon}_\heartsuit$$

Suppose we would like to translate the axiom $\heartsuit(\spadesuit p \to \heartsuit q) \to (\spadesuit p \to \heartsuit q)$ into a rule equivalent to it over $\mathcal{R}$. The axiom can be seen as a substitution instance of the non-iterative axiom $\heartsuit(r \to s) \to (r \to s)$ under the substitution σ with $\sigma(r) = \spadesuit p$ and $\sigma(s) = \heartsuit p$. Translating the latter we first get the ground sequent $\overline{\Rightarrow \heartsuit(r \to s) \to (r \to s)}$, and resolving propositional logic on the top level of the formula then yields $\overline{\heartsuit(r \to s), r \Rightarrow s}$. Since our rule format does not allow nested connectives in the conclusion we now introduce a fresh variable t for the formula $r \to s$ and premisses $t \Rightarrow (r \to s)$ and $(r \to s) \Rightarrow t$ which ensure that t and $r \to s$ are equivalent. As we will see due to the monotonicity of $\heartsuit$ we actually only need the premiss $t \Rightarrow (r \to s)$, which gives us

$$\frac{t \Rightarrow r \to s}{\heartsuit t, r \Rightarrow s} \ .$$

Now resolving the propositional connective in the premiss and substituting the original formulae $\spadesuit p$ for r and $\heartsuit q$ for s gives us

$$\frac{t, \spadesuit p \Rightarrow \heartsuit q}{\heartsuit t, \spadesuit p \Rightarrow \heartsuit q} \ .$$

Finally, using the facts that the modality $\spadesuit$ is normal and that the modality $\heartsuit$ is serial we may turn the variables p and q above into the context restriction $\mathcal{C} := \langle \{\spadesuit p\}, \{\heartsuit q\} \rangle$, giving the rule $\{(s \Rightarrow\ ; \mathcal{C})\}/\heartsuit s \Rightarrow$.

There are some peculiarities to the procedure described above. The first is that all the propositional logic occurring on the top level of the formula or under the modalities must be resolvable, that is the respective rules for the propositional connectives must be invertible. While for classical propositional logic as the underlying propositional logic this is always the case, for intuitionistic propositional logic we need to be more careful. The second is that we can only turn the formulae $\spadesuit p$ and $\heartsuit q$ above into context restrictions, because all their occurrences are on the *same* side of the sequents in the premisses and the conclusion, and because they occur exactly once in the conclusion and at least once in the premisses. In order to express these requirements formally we introduce the notion of a *resolvable formula*, where we furthermore make the subformulae occurring only negatively or positively explicit in the sets C_ℓ resp. C_r:

Definition 3.2.2 (ci)**.** Let C_ℓ, C and C_r be sets of formulae. We simultaneously define the sets $\mathcal{F}_r{}^{\mathsf{c}}(C_\ell, C, C_r)$ of *classically right resolvable formulae* and $\mathcal{F}_\ell{}^{\mathsf{c}}(C_\ell, C, C_r)$ of *classically left resolvable formulae* for the triple (C_ℓ, C, C_r) as well as their intuitionistic versions $\mathcal{F}_r{}^{\mathsf{i}}(C_\ell, C, C_r)$ and $\mathcal{F}_\ell{}^{\mathsf{i}}(C_\ell, C, C_r)$ by the following grammar, where for $\mathcal{F}_r{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$ the starting variable is $P_r^{[\mathsf{ci}]}$ and for $\mathcal{F}_\ell{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$ it is $P_\ell^{[\mathsf{ci}]}$.

$$\begin{array}{ll} P_r^{[\mathsf{ci}]} ::= P_r^{[\mathsf{c}]} \lor P_r^{[\mathsf{c}]} \mid P_r^{[\mathsf{ci}]} \land P_r^{[\mathsf{ci}]} \mid P_\ell^{[\mathsf{ci}]} \to P_r^{[\mathsf{ci}]} \mid A_r \mid B \mid \bot \mid \top & \text{where } A_r \in C_r, B \in C \\ P_\ell^{[\mathsf{ci}]} ::= P_\ell^{[\mathsf{ci}]} \lor P_\ell^{[\mathsf{ci}]} \mid P_\ell^{[\mathsf{ci}]} \land P_\ell^{[\mathsf{ci}]} \mid P_r^{[\mathsf{c}]} \to P_\ell^{[\mathsf{c}]} \mid A_\ell \mid B \mid \bot \mid \top & \text{where } A_\ell \in C_\ell, B \in C \ . \end{array}$$

Remark 3.2.3. It is obvious from the definition that if a formula is left (or right) resolvable for $(C_\ell \cup \{A\}, C, C_r \cup \{B\})$, then it is also left (resp. right) resolvable for $(C_\ell, C \cup \{A\}, C_r \cup \{B\})$ and for $(C_\ell \cup \{A\}, C \cup \{B\}, C_r)$.

Example 3.2.4.
1. The formula $\heartsuit p \land ((\heartsuit p \lor q) \to r)$ is intuitionistically and classically right resolvable for $(\{q\}, \{\heartsuit p\}, \{r\})$. Furthermore, it is classically left resolvable for $(\{r\}, \{\heartsuit p\}, \{q\})$ but it is not intuitionistically left resolvable for any triple.

2. Both $p \vee q$ and $(p \to q) \to \bot$ are classically right resolvable for $(\emptyset, \emptyset, \{p, q\})$ resp. $(\{q\}, \emptyset, \{p\})$. Both formulae are not intuitionistically right resolvable for any triple.

Intuitively for a formula A in $\mathcal{F}_r{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$ the sequent $\Rightarrow A$ can be broken up or resolved using invertibility of the propositional rules in such a way that in the resulting set of sequents formulae in C_ℓ (resp. C_r) occur only on the left hand side (resp. right hand side), whereas formulae in C might occur on both sides. We capture this in the following lemma.

Lemma 3.2.5 (ci)**.** *Let C_ℓ, C and C_r be sets of formulae and let $\Gamma \Rightarrow \Delta$ be a sequent such that every formula in Γ is in $\mathcal{F}_\ell{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$ and every formula in Δ is in $\mathcal{F}_r{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$. Then there is a set $S = \{\Gamma_i \Rightarrow \Delta_i \mid i \in I\}$ of sequents such that*

1. $S \vdash_{\mathsf{G[ci]CutConW}} \Gamma \Rightarrow \Delta$
2. $\Gamma \Rightarrow \Delta \vdash_{\mathsf{G[ci]CutConW}} \Gamma_i \Rightarrow \Delta_i$ *for every* $i \in I$
3. *for every* $i \in I$*: every formula in* Γ_i *is in* $C_\ell \cup C$
4. *for every* $i \in I$*: every formula in* Δ_i *is in* $C_r \cup C$.

Proof. The proof is by simultaneous induction on the structure of the formulae in $\mathcal{F}_\ell{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$ and $\mathcal{F}_r{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$ and essentially works by inverting the propositional rules using the rules in $\mathsf{G[ci]CutCon}$ until we arrive at a set of sequents with the properties stated above.

As an example for the induction step in the classical case, suppose we have a sequent $\Gamma, A \Rightarrow \Delta$ such that $A \in \mathcal{F}_\ell{}^{\mathsf{c}}(C_\ell, C, C_r)$ with $A = A_1 \to A_2$ for $A_1 \in \mathcal{F}_r{}^{\mathsf{c}}(C_\ell, C, C_r)$ and $A_2 \in \mathcal{F}_\ell{}^{\mathsf{c}}(C_\ell, C, C_r)$, and such that $\mathsf{Supp}\,(\Gamma) \subseteq \mathcal{F}_\ell{}^{\mathsf{c}}(C_\ell, C, C_r)$ and $\mathsf{Supp}\,(\Delta) \subseteq \mathcal{F}_r{}^{\mathsf{c}}(C_\ell, C, C_r)$. Then $\Gamma, A \Rightarrow \Delta$ is equivalent to the two sequents $\Gamma \Rightarrow A_1, \Delta$ and $\Gamma, A_2 \Rightarrow \Delta$ as seen by an application of $\to_{\mathsf{L}}$ in the one direction and the derivations

$$\dfrac{\dfrac{\dfrac{}{A_1 \Rightarrow A_1, A_2}\,\mathcal{A}}{\Rightarrow A_1, A_1 \to A_2}\,{\to_{\mathsf{R}}} \qquad \Gamma, A_1 \to A_2 \Rightarrow \Delta}{\Gamma \Rightarrow A_1, \Delta}\,\mathsf{Cut} \qquad \dfrac{\dfrac{\dfrac{}{A_1, A_2 \Rightarrow A_2}\,\mathcal{A}}{A_2 \Rightarrow A_1 \to A_2}\,{\to_{\mathsf{R}}} \qquad \Gamma, A_1 \to A_2 \Rightarrow \Delta}{\Gamma, A_2 \Rightarrow \Delta}\,\mathsf{Cut}$$

in the other direction. The cases for the other binary boolean connectives are analogous. In case A is $\bot$ we use the rule $\bot_{\mathsf{L}}$ and Weakening.

In the intuitionistic case the definition of right resp. left resolvable formulae ensures that we do not need to invert the problematic rules $\vee_{\mathsf{R}}$ or $\to_{\mathsf{L}}$. The remaining cases are analogous to the classical situation. The only slightly interesting case is that of a sequent $\Gamma \Rightarrow A$ with $\mathsf{Supp}\,(\Gamma) \subseteq \mathcal{F}_\ell{}^{\mathsf{i}}(C_\ell, C, C_r)$ and $A = A_1 \to A_2$. In this case we have

$$\dfrac{\Gamma \Rightarrow A \to B \qquad \dfrac{\dfrac{}{B, A \Rightarrow B}\,\mathcal{A} \qquad \dfrac{}{A \Rightarrow A}\,\mathcal{A}}{A \to B, A \Rightarrow B}\,{\to_{\mathsf{L}}}}{\Gamma, A \Rightarrow B}\,\mathsf{Cut}$$

and on the other hand the sequent $\Gamma \Rightarrow A \to B$ follows directly from $\Gamma, A \Rightarrow B$ using $\to_{\mathsf{R}}$. $\square$

The set of sequents guaranteed by the previous lemma need not be unique. E.g., if for a formula A we have $\bot \to A \in C_r$ and $A \in C_r$, then the sequent $\Rightarrow \bot \to A$ already is in the form stated in the lemma, but it is also equivalent to the sequent $\bot \Rightarrow A$. However, this will not be problematic. Using the previous Lemma it is then possible to break up formulae in $\mathcal{F}_r{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$ and to turn them into a normal form:

Definition 3.2.6 (ci)**.** For a sequent $\Gamma \Rightarrow \Delta$ with $\mathsf{Supp}\,(\Gamma) \subseteq \mathcal{F}_\ell{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$ and $\mathsf{Supp}\,(\Delta) \subseteq \mathcal{F}_r{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$ a set of sequents resulting from its deconstruction according to Lemma 3.2.5 is called a (C_ℓ, C, C_r)-*normal form* of $\Gamma \Rightarrow \Delta$. A (C_ℓ, C, C_r)-*normal form* of a formula $A \in \mathcal{F}_r{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$ is a (C_ℓ, C, C_r)-normal form of the sequent $\Rightarrow A$.

To compute a (C_ℓ, C, C_r)-normal form of a formula A from $\mathcal{F}_r{}^{[\mathsf{ci}]}(C_\ell, C, C_r)$ it is thus enough to attempt backwards proof search using the propositional rules for the sequent $\Rightarrow A$ until the sequents at the leafs of the derivation have only formulae from $C_\ell \cup C$ on their left hand side and only formulae from $C_r \cup C$ on their right hand side, and then take the set of the sequents at the leafs of the derivation.

Example 3.2.7.

1. If we attempt backwards proof search using the rules in $\mathsf{G[cim]}$ for the formula $\heartsuit p \wedge ((\heartsuit p \vee q) \to r)$ from Example 3.2.4,1 we get the following:

$$\dfrac{\Rightarrow \heartsuit p \qquad \dfrac{\dfrac{\heartsuit p \Rightarrow r \qquad q \Rightarrow r}{\heartsuit p \vee q \Rightarrow r}\,\vee_{\mathsf{L}}}{\Rightarrow (\heartsuit p \vee q) \to r}\,\to_{\mathsf{R}}}{\Rightarrow \heartsuit p \wedge ((\heartsuit p \vee q) \to r)}\,\wedge_{\mathsf{R}}$$

   The formula $\heartsuit p$ occurs both on the left and on the right hand side of the sequents $\Rightarrow \heartsuit p$ and $\heartsuit p, q \Rightarrow r$ at the leafs of the attempted derivation, whereas q and r occur only on the left (resp. right) hand side. This gives a classical and intuitionistic $(\{q\}, \{\heartsuit p\}, \{r\})$-normal form $\{\Rightarrow \heartsuit p,\ \heartsuit p \Rightarrow r,\ q \Rightarrow r\}$ for the formula $\heartsuit p \wedge ((\heartsuit p \vee q) \to r)$.

2. For the formula $(p \to q) \to \bot$ we get the classical derivation

$$\dfrac{\dfrac{q \Rightarrow \bot \qquad \Rightarrow p, \bot}{p \to q \Rightarrow \bot}\,\to_{\mathsf{L}}}{\Rightarrow (p \to q) \to \bot}\,\to_{\mathsf{R}}$$

   and thus a classical $(\{q\}, \emptyset, \{p\})$-normal form $\{q \Rightarrow \bot,\ \Rightarrow p, \bot\}$ for $(p \to q) \to \bot$.

Remark 3.2.8. If A is a (classically or intuitionistically) right resolvable propositional formula for $(\emptyset, \mathsf{Var}, \emptyset)$, then by permutability of the propositional rules its $(\emptyset, \mathsf{Var}, \emptyset)$-normal form is unique. Moreover, Definition 3.2.6 yields the *regular normal form* of A in the sense of [NvP01, p.128] if we consider the conjunction of the trace formulae of the sequents in its $(\emptyset, \mathsf{Var}, \emptyset)$-normal form.

The notion of right or left resolvable formulae now enables us to formally define a class of translatable formulae. For the sake of readability we restrict ourselves to unary and monotone modalities, i.e., we assume presence of the rules in $\mathsf{Mon} := \{\mathsf{Mon}_\heartsuit \mid \heartsuit \in \Lambda\}$.

Definition 3.2.9 (ci)**.** Suppose that Λ is a set of monotone and unary modalities. Let C_ℓ and C_r be sets of formulae with variables in W_ℓ and W_r respectively, and let V be a set of variables such that for all formulae $A, B \in C_\ell \cup C_r$ we have $\mathsf{var}(A) \cap \mathsf{var}(B) = \emptyset$ and $\mathsf{var}(A) \cap V = \emptyset$. The set of *intuitionistically (resp. classically) translatable clauses with context formulae in* (C_ℓ, C_r) *and variables in* V is given by the following grammar (with initial variable $S^{[\mathsf{ci}]}$)

$$
\begin{aligned}
S^{[\mathsf{ci}]} &::= L^{[\mathsf{ci}]} \to R^{[\mathsf{ci}]}\\
L^{[\mathsf{ci}]} &::= L^{[\mathsf{ci}]} \wedge L^{[\mathsf{ci}]} \mid \heartsuit P_r \mid A_\ell \mid p \mid \bot \mid \top \quad \text{where } \heartsuit \in \Lambda, P_r \in \mathcal{F}_r{}^{[\mathsf{ci}]}(C_\ell, V, C_r), A_\ell \in C_\ell, p \in V\\
R^{[\mathsf{ci}]} &::= R^{[\mathsf{c}]} \vee R^{[\mathsf{c}]} \mid \heartsuit P_\ell \mid A_r \mid p \mid \bot \mid \top \quad \text{where } \heartsuit \in \Lambda, P_\ell \in \mathcal{F}_\ell{}^{[\mathsf{ci}]}(C_\ell, V, C_r), A_r \in C_r, p \in V
\end{aligned}
$$

together with the global restriction that every formula in $C_\ell \cup C_r$ occurs at most once on the top level of the clause (i.e. not in the scope of a modality), and occurs on the top level if and only if it occurs under a modality. We also call the formulae in $C_\ell \cup C_r$ the *context formulae* of such a clause. A clause is *intuitionistically (resp. classically) translatable* if there is a triple (C_ℓ, V, C_r) such that the clause is intuitionistically (resp. classically) translatable with context formulae in (C_ℓ, C_r) and variables in V.

In other words, a classically translatable clause with context formulae in (C_ℓ, C_r) and variables in V is simply a clause $\bigwedge_{i=1}^n A_i \to \bigvee_{j=1}^m B_j$ where for $i \leq n$ we have $A_i \in \Lambda(\mathcal{F}_r{}^{[\mathsf{c}]}(C_\ell, V, C_r)) \cup C_\ell \cup V \cup \{\bot, \top\}$ and for $j \leq m$ we have $B_j \in \Lambda(\mathcal{F}_\ell{}^{[\mathsf{c}]}(C_\ell, V, C_r)) \cup C_r \cup V \cup \{\bot, \top\}$ satisfying the conditions on occurrences of formulae in $C_\ell \cup C_r$. In the intuitionistic case the situation is similar with $m \in \{0, 1\}$. Thus a translatable clause is essentially a non-iterative formula where some of the variables have been substituted with arbitrary modal formulae which appear positively (resp. negatively) on the top level of the formula if and only if they appear in the same polarity under a modality.

Example 3.2.10.

1. The axiom $(\mathsf{4}\Box) = \Box s \to \Box\Box s$ is an intuitionistically and classically translatable clause with context formulae in $(\{\Box s\}, \emptyset)$.

2. The axiom $(\mathsf{5}) = \Diamond s \to \Box \Diamond s$ is an intuitionistically and classically translatable clause with context formulae in $(\{\Diamond s\}, \emptyset)$. The version $(\mathsf{5}\Box) = \neg\Box s \to \Box\neg\Box s$ is a classically translatable clause where the context formulae can be taken to be in $(\{\neg\Box s\}, \emptyset)$ or in $(\emptyset, \{\Box s\})$.

3. The axiom $(\mathsf{T}\Box) = \Box s \to s$ is an intuitionistically and classically translatable clause where the context formulae can be taken to be either in $(\emptyset, \emptyset)$ or in $(\emptyset, \{s\})$.

4. The axiom $(\mathsf{IK2}) = \Box(s \to t) \wedge \Diamond s \to \Diamond t$ is an intuitionistically and classically translatable clause with context formulae in $(\emptyset, \emptyset)$.

5. The axiom $(\mathsf{IK4}) = \Diamond(p \vee q) \to \Diamond p \vee \Diamond q$ is a classically translatable clause with context formulae $(\emptyset, \emptyset)$. It is not an intuitionistically translatable clause.

6. The axiom $(\mathsf{B}) = p \to \Box\neg\Box\neg p$ is neither a classically nor an intuitionistically translatable clause, since the variable p occurs both on the top level of the formula (and thus would need to be a context formula or a variable in V) and under two modalities (and thus would need to be a proper subformula of a context formula).

7. The axiom $(\mathsf{L}) = \Box(\Box p \to p) \to \Box p$ is neither a classically nor an intuitionistically translatable clause, since the variable p occurs under two modalities and thus cannot be a variable in V, and moreover the formulae $\Box p \to p$ and p do not occur on the top level of the formula and thus cannot be context formulae.

Remark 3.2.11. For the sake of readability we only considered unary and monotone modalities. It is not too difficult to generalise Definition 3.2.9 to n-ary and not necessarily monotone modalities, though. Let us say that for a rule set $\mathcal{R}$ and $i \leq n$ an n-ary modality $\heartsuit$ is *monotone in the i-th argument for $\mathcal{R}$* if the rule

$$\begin{aligned}\mathsf{Mon}_{\heartsuit,i} := &\{(p_i \Rightarrow q_i; \mathcal{C}_\emptyset)\} \cup \{(p_j \Rightarrow q_j; \mathcal{C}_\emptyset), (q_j \Rightarrow p_j; \mathcal{C}_\emptyset) \mid j \leq n,\ j \neq i\} \,/ \\ &\heartsuit(p_1, \dots, p_n) \Rightarrow \heartsuit(q_1, \dots, q_n)\end{aligned}$$

is derivable in $\mathcal{R}$. Similarly, we say that $\heartsuit$ is *antitone in the i-th argument for $\mathcal{R}$* if the rule

$$\begin{aligned}\mathsf{Ant}_{\heartsuit,i} := &\{(q_i \Rightarrow p_i; \mathcal{C}_\emptyset)\} \cup \{(p_j \Rightarrow q_j; \mathcal{C}_\emptyset), (q_j \Rightarrow p_j; \mathcal{C}_\emptyset) \mid j \leq n,\ j \neq i\} \,/ \\ &\heartsuit(p_1, \dots, p_n) \Rightarrow \heartsuit(q_1, \dots, q_n)\end{aligned}$$

is derivable in $\mathcal{R}$. Then in the definition of a translatable clause for the variable $L^{[\mathsf{ci}]}$ we simply replace the entry $\heartsuit P_r$ by $\heartsuit(P_1, \dots, P_n)$ where $\heartsuit$ is an n-ary modality from Λ and for $1 \leq i \leq n$ we have: $P_i \in {\mathcal{F}_r}^{[\mathsf{ci}]}(C_\ell, V, C_r)$ if $\heartsuit$ is monotone in the i-th argument for $\mathcal{R}$; $P_i \in {\mathcal{F}_\ell}^{[\mathsf{ci}]}(C_\ell, V, C_r)$ if $\heartsuit$ is antitone in the i-th argument for $\mathcal{R}$; and $P_i \in {\mathcal{F}_r}^{[\mathsf{ci}]}(C_\ell, V, C_r) \cap {\mathcal{F}_\ell}^{[\mathsf{ci}]}(C_\ell, V, C_r)$ otherwise. For $R^{[\mathsf{ci}]}$ the entry $\heartsuit P_\ell$ is modified similarly.

The context formulae of a translatable clause will of course correspond to the context formulae in the resulting rule. But since the number of context formulae in a translatable clause is fixed, in the result of the translation we have a fixed number of context formulae as well. We capture this in the notion of a proto rule for a rule with restrictions.

Definition 3.2.12. Given a rule with restrictions $R = \mathcal{P}/\Sigma \Rightarrow \Pi$ a *proto rule for R* is a tuple $(R; \Gamma \Rightarrow \Delta)$ given by a context $\Gamma \Rightarrow \Delta \in \mathcal{S}(\mathcal{F})$ such that

1. no propositional variable occurs more than once in $\Gamma \Rightarrow \Delta$
2. no propositional variable occurs both in $\Gamma \Rightarrow \Delta$ and R
3. if $\Gamma \Rightarrow \Delta \neq \ \Rightarrow$, then $(\Gamma \Rightarrow \Delta) \restriction_{\mathcal{C}} \neq \ \Rightarrow$ for every restriction $\mathcal{C}$ of R.

We often leave the context implicit and write $\widehat{R}$ for a proto rule for R. An *application* of a proto rule $\widehat{R} = (R; \Gamma \Rightarrow \Delta)$ is given by a substitution $\sigma : \mathsf{Var} \to \mathcal{F}$ and a context $\Theta \Rightarrow \Xi$ where $(\Theta \Rightarrow \Xi) \restriction_{\mathcal{C}} = \ \Rightarrow$ for every restriction $\mathcal{C}$ of R, and is the same as the application of R with substitution σ and context $\Gamma\sigma, \Theta \Rightarrow \Delta\sigma, \Xi$ according to Definition 2.3.3. Derivability using proto rules and notions of equivalence are defined as expected.

Informally, the difference between rules and proto rules is that in proto rules the premisses *including the context* are fixed up to substitution, while in rules also the *number* of the context formulae in the premisses may vary. Clause 3 in the above definition ensures that the context $\Gamma \Rightarrow \Delta$ only contains context formulae which are copied into at least one of the premisses. Together with the condition that none of the formulae in a context $\Theta \Rightarrow \Xi$ for an *application* of the proto rule are copied into the premiss this means that a proto rule for a rule really specifies exactly the number of context formulae which are copied into the premisses. In the notation of Lemma 2.4.4 a proto rule $\widehat{R}$ for a rule $R = \mathcal{P}/\Sigma \Rightarrow \Pi$ given by a context $\Gamma \Rightarrow \Delta$ has premisses $\mathcal{P}(\Gamma \Rightarrow \Delta)$ and conclusion $\Gamma, \Sigma \Rightarrow \Pi, \Delta$.

Example 3.2.13.

1. Consider the rule $\mathsf{K}_n = \{(p_1, \dots, p_n \Rightarrow q; \mathcal{C}_\emptyset)\}/\Box p_1, \dots, \Box p_n \Rightarrow \Box q$. Since the only restriction occurring in it is $\mathcal{C}_\emptyset = \langle \emptyset, \emptyset \rangle$, the only proto rule for K_n is given by the empty context $\Rightarrow$. An application of this proto rule is given by a substitution and an arbitrary context.

2. For the rule $R_{\mathsf{T}} = \{(p \Rightarrow ; \mathcal{C}_{\mathsf{id}})\}/\Box p \Rightarrow$ proto rules are given by arbitrary contexts (satisfying the conditions on the variables), since every such sequent satisfies the restriction $\mathcal{C}_{\mathsf{id}}$. Applications for these proto rules are given by a substitution and the empty context. Thus e.g. the proto rule given by the sequent $q_1, q_2 \Rightarrow r$ has applications
$$\frac{A_1, A_2, B \Rightarrow C}{A_1, A_2, \Box B \Rightarrow C}$$
for arbitrary formulae A_1, A_2, B, C.

3. Proto rules for the rule $\mathsf{4}_0 = \{(\Rightarrow q; \mathcal{C}_4)\}/ \Rightarrow \Box q$ are given by sequents $\Box A_1, \dots, \Box A_n \Rightarrow$ for a fixed $n \in \mathbb{N}$ (satisfying the conditions on the variables). Applications of these proto rules are given by a substitution and a context consisting on the left only of formulae whose main connective is not $\Box$. Thus e.g. the proto rules given by the contexts $\Box p_1, \Box p_2 \Rightarrow$ and $\Rightarrow$ respectively have applications
$$\frac{\Box A_1, \Box A_2 \Rightarrow B}{\Gamma, \Box A_1, \Box A_2 \Rightarrow \Box B, \Delta} \qquad \text{and} \qquad \frac{\Rightarrow B}{\Gamma \Rightarrow \Box B, \Delta}$$

respectively, where no formula in Γ has $\Box$ as its main connective (and in the intuitionistic case Δ is empty).

Now the procedure sketched in the beginning of this section gives us a translation.

Theorem 3.2.14 (ci)**.** *Let Λ be a set of unary modalities. Every intuitionistically resp. classically translatable clause A is equivalent over* Mon *to a proto rule. Moreover, there is an automatic procedure that given an axiom produces such an equivalent proto rule.*

For the sake of presentation in the context of the proof of this statement we will take a more relaxed attitude concerning the notion of sequent rules and will take the premisses of a rule to be sets of arbitrary sequents and the conclusion to be an arbitrary sequent. This allows us e.g. as an intermediate step to move complex formulae from the conclusion into the premisses, the result of which is not a rule with context restrictions. Of course the final result will be a rule with context restrictions again. Connected notions such as applications of rules or derivations are adjusted as expected. We will make use of the following two easy lemmata, the first of which is also called the Ackermann Lemma in [CGT08] and allows us to move propositional variables from the conclusion into the premisses.

Lemma 3.2.15 (Folklore)(ci)**.** *For $n, m \geq 0$ and formulae $A_1, \dots, A_n, B_1, \dots, B_m$ the following two rules are equivalent:*

$$\frac{\mathcal{P}}{A_1, \dots, A_n, \Gamma \Rightarrow \Delta, B_1, \dots, B_m} \qquad\qquad \frac{\mathcal{P} \,\cup\, \{\Sigma \Rightarrow A_i, \Pi \mid i \leq n\} \,\cup\, \{\Sigma, B_j \Rightarrow \Pi \mid j \leq m\}}{\Gamma, \Sigma \Rightarrow \Delta, \Pi}$$

In the asymmetric setting for $n > 0$ we have $\Pi = \emptyset$ and for $m > 0$ we have $\Delta = \emptyset$.

Proof. Using Cut and the fact that sequents $A_1, \dots, A_n \Rightarrow B_1, \dots B_m, A_i$ and $A_1, \dots, A_n, B_j \Rightarrow B_1, \dots, B_m$ are derivable using the axiom rule for $1 \leq i \leq n$ and $1 \leq j \leq m$. $\Box$

The second lemma we will use states that we can modify the rules so that the principal formulae consist only of modalised variables.

Lemma 3.2.16 (ci)**.** *For every unary monotone modality $\heartsuit$ and formula D the rules*

$$\frac{\mathcal{P}}{\Gamma, \heartsuit D \Rightarrow \Delta} \qquad \textit{resp.} \qquad \frac{\mathcal{P}}{\Gamma \Rightarrow \heartsuit D, \Delta}$$

are equivalent over $\mathsf{Mon}_\heartsuit$ *to the rules*

$$\frac{\mathcal{P} \,\cup\, \{s_D \Rightarrow D\}}{\Gamma, \heartsuit s_D \Rightarrow \Delta} \qquad \textit{resp.} \qquad \frac{\mathcal{P} \,\cup\, \{D \Rightarrow s_D\}}{\Gamma \Rightarrow \heartsuit s_D, \Delta}$$

respectively, where s_D is a fresh propositional variable.

Proof. Using the monotonicity rule $\mathsf{Mon}_\heartsuit$ and Cut and the fact that the sequent $D \Rightarrow D$ is derivable using the axiom rule. $\Box$

Proof of Theorem 3.2.14. Suppose that A is a translatable clause with context formulae in (C_ℓ, C_r) and variables in V. We assume that the variables in V are ordered in some way. W.l.o.g. we furthermore assume that C_ℓ and C_r do not contain single variables (this can always be achieved by transferring them into V, see Remark 3.2.3), and that every formula in C_ℓ and C_r occurs in A. Then A is of the form

$$\bigwedge_{\heartsuit_{P_\ell} P_\ell \in F_\ell} \heartsuit_{P_\ell} P_\ell \wedge \bigwedge_{C \in C_\ell} C \wedge \bigwedge_{q \in Q_\ell} q \to \bigvee_{\heartsuit_{P_r} P_r \in F_r} \heartsuit_{P_r} P_r \vee \bigvee_{D \in C_r} D \vee \bigvee_{r \in Q_r} r$$

where the formulae P_ℓ are in $\mathcal{F}_r{}^{[\mathsf{ci}]}(C_\ell, V, C_r)$, the formulae P_r are in $\mathcal{F}_\ell{}^{[\mathsf{ci}]}(C_\ell, V, C_r)$, and we have $Q_\ell \subseteq V$ and $Q_r \subseteq V$. The first step is to turn this into a ground sequent $\overline{\Rightarrow A}$, which by Lemma 3.2.5 is equivalent to the following (where we slightly abuse notation and write e.g. F_ℓ for the multiset consisting of all the formulae in the set F_ℓ)

$$\frac{}{F_\ell, C_\ell, Q_\ell \Rightarrow F_r, C_r, Q_r} \ .$$

Now using Lemma 3.2.16 we introduce fresh variables for the formulae P_ℓ resp. P_r occurring under the modalities and premisses ensuring that the variables are equivalent to the original formulae. This yields

$$\frac{\{s_{P_\ell} \Rightarrow P_\ell \mid \heartsuit_{P_\ell} P_\ell \in Q_\ell\} \ \cup \ \{P_r \Rightarrow s_{P_r} \mid \heartsuit_{P_r} P_r \in Q_r\}}{\{\heartsuit_{P_\ell} s_{P_\ell} \mid P_\ell \in F_\ell\}, C_\ell, Q_\ell \Rightarrow \{\heartsuit_{P_r} s_{P_r} \mid P_r \in F_r\}, C_r, Q_r} \ .$$

Since the formulae P_ℓ (resp. P_r) are in $\mathcal{F}_r{}^{[\mathsf{ci}]}(C_\ell, V, C_r)$ (resp. $\mathcal{F}_\ell{}^{[\mathsf{ci}]}(C_\ell, V, C_r)$) we may now equivalently replace these new premisses by (C_ℓ, V, C_r)-normal forms using Lemma 3.2.5. In particular formulae in C_ℓ (resp. C_r) only occur on the left (resp. right) hand side of the resulting premisses. Moreover, since by Definition 3.2.9 every formula in $C_\ell \cup C_r$ occurs on the top level of the axiom if and only if it occurs under a modality we get that every of these formulae occurs in the conclusion of the rule if and only if it occurs in at least one premiss. Let us call this set of premisses $\mathcal{P}$. The next step is to use Lemma 3.2.15 to move all the variables in $Q_\ell \cup Q_r$ from the conclusion into the premisses. This yields

$$\frac{\mathcal{P} \ \cup \ \{t_\ell \Rightarrow q, t_r \mid q \in Q_\ell\} \ \cup \ \{t_\ell, r \Rightarrow t_r \mid r \in Q_r\}}{\{\heartsuit_{P_\ell} s_{P_\ell} \mid P_\ell \in F_\ell\}, C_\ell, t_\ell \Rightarrow \{\heartsuit_{P_r} s_{P_r} \mid P_r \in F_r\}, C_r, t_r}$$

where the new context is replaced with fresh propositional variables t_ℓ and t_r. The final step is to apply the obvious adaption of the process of variable elimination (Definition 2.4.1) to the premisses to eliminate all variables in V from the premisses by performing all possible cuts between premisses on these variables. We assume that this is done in the order given by the ordering on V. This ensures that a variable occurs in the premisses only if it occurs in the conclusion. The resulting rule is seen to be equivalent over Mon to the rule given above using the methods of the proof of Lemma 2.4.5. Moreover, using Weakening and the rules in $\mathsf{G[ci]CutCon}$, it is equivalent to a proto rule for the rule with restrictions where we replace in every premiss all occurring context formulae, say $C_1, \ldots, C_n \Rightarrow D_1, \ldots, D_m$, by the corresponding context restriction $\langle \{C_1, \ldots, C_n\}, \{D_1, \ldots, D_m\} \rangle$, and similarly the variables t_ℓ, t_r by the restriction $\langle \{t_\ell\}, \{t_r\} \rangle$. $\square$

Example 3.2.17. Continuing Example 3.2.10 above we have

1. The intuitionistically and classically translatable clause $\Box p \wedge \Box q \to \Box(p \wedge q)$ with context formulae in $(\emptyset, \emptyset)$ and variables in $\{p, q\}$ is transformed in a first step into $\overline{\Box p, \Box q \Rightarrow \Box(p \wedge q)}$. In the next step the formulae p, q and $p \wedge q$ under the modalities are replaced by fresh variables s_p, s_q and $s_{p \wedge q}$ and we introduce new premisses to obtain

$$\frac{s_p \Rightarrow p \qquad s_q \Rightarrow q \qquad p, q \Rightarrow s_{p \wedge q}}{\Box s_p, \Box s_q \Rightarrow \Box s_{p \wedge q}} \, .$$

   Since there are no variables in the conclusion, the next step can be omitted, leaving the final step of variable elimination on the variables p and q. This gives the well-known rule

$$\frac{s_p, s_q \Rightarrow s_{p \wedge q}}{\Box s_p, \Box s_q \Rightarrow \Box s_{p \wedge q}} \; \mathsf{K}_2 \, .$$

2. Similarly, the intuitionistically and classically translatable clause $(\mathsf{IK2}) = \Box(p \to q) \wedge \Diamond p \to \Diamond q$ is translated into the rule

$$\frac{s_{p \to q}, s_p \Rightarrow s_q}{\Box s_{p \to q}, \Diamond s_p \Rightarrow \Diamond s_q} \, .$$

3. The intuitionistically and classically translatable clause $(\mathsf{4}\Box) = \Box s \to \Box\Box s$ with context formulae in $(\{\Box s\}, \emptyset)$ is first transformed into $\overline{\Box s \Rightarrow \Box\Box s}$. Then the formula $\Box s$ under the modality is replaced by a fresh variable $s_{\Box s}$ to give

$$\frac{\Box s \Rightarrow s_{\Box s}}{\Box s \Rightarrow \Box s_{\Box s}} \, ,$$

   which as seen in Example 3.2.13 is a proto rule for the rule $\mathsf{4}_0$.

4. The classically translatable clause $(\mathsf{5}\Box) = \neg\Box s \to \Box\neg\Box s$ with context formulae in $(\emptyset, \{\Box s\})$ is similarly transformed into the proto rule

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

   for the rule R_{45}.

5. The intuitionistically and classically translatable clause $(\mathsf{T}\Box) = \Box s \to s$ with context formulae in $(\emptyset, \emptyset)$ is first transformed into $\overline{\Box s \Rightarrow s}$. Then the variable s on the right

hand side is moved into the premisses using Lemma 3.2.15, which gives the proto rule

$$\frac{p_\ell, s \Rightarrow p_r}{p_\ell, \Box s \Rightarrow p_r}$$

for the rule R_{T}. If the context formulae are taken to be in $(\emptyset, \{s\})$, then we obtain the proto rule

$$\frac{p \Rightarrow s}{\Box p \Rightarrow s}\,.$$

Remark 3.2.18. While we stated Theorem 3.2.14 only for unary and monotone modalities, it is straightforwardly adapted to modalities of higher arities as follows. If a formula $\heartsuit(P_1, \ldots, P_n)$ occurs on the left side of the implication in a clause A which is translatable according to Remark 3.2.11, then we introduce fresh variables $s_{P_1}, \ldots, s_{P_n}$ and for $1 \leq i \leq n$ premisses

$s_{P_i} \Rightarrow P_i$	if $\heartsuit$ is monotone in the ith argument
$P_i \Rightarrow s_{P_i}$	if $\heartsuit$ is antitone in the ith argument
$s_{P_i} \Rightarrow P_i$ and $P_i \Rightarrow s_{P_i}$	otherwise.

The modified definition of a translated clause then ensures that when we replace these new premisses by (C_ℓ, V, C_r)-normal forms the context formulae behave in the correct way. In particular, instead of the rule Mon we use the derivation of the rule $\mathsf{Mon}_{\heartsuit,i}$ resp. $\mathsf{Ant}_{\heartsuit,i}$ to show that in case the modality is monotone resp. antitone in the i-th argument it is enough to include only the sequent $s_{P_i} \Rightarrow P_i$ resp. $P_i \Rightarrow s_{P_i}$ in the premisses. Note that in order to do this we need that the rule $\mathsf{Mon}_{\heartsuit,i}$ resp. $\mathsf{Ant}_{\heartsuit,i}$ is derivable and not just admissible in the rule set. On the other hand this means that we do not need to explicitly assume that these rules are in the base rule set.

Unfortunately, Theorem 3.2.14 only tells us which axioms we can translate into *proto rules* for rules with restrictions. But of course we would like to know which axioms we can translate into rules with restrictions themselves. The main idea here is that if the context formulae of a rule with restrictions absorb conjunctions (resp. disjunctions) in the right way, then we can restrict ourselves to proto rules where every context formula occurs exactly once.

Definition 3.2.19 (ci)**.** A formula A with free variables $p_1, \ldots, p_n = \vec{p}$ is *intuitionistically (resp. classically) left normal* for a set $\mathcal{R}$ of rules if for every $k \geq 0$ there are formulae $B_1, \ldots, B_n = \vec{B}$ such that $\vdash_{\mathsf{G[ci]CutCon}\mathcal{R}} \Rightarrow \bigwedge_{i=1}^k A^i \leftrightarrow A\sigma_{\vec{p}}^{\vec{B}}$ where A^i is the result of injectively renaming the propositional variables $p_1, \ldots, p_n$ in A to fresh variables $p_1^i, \ldots, p_n^i$ and $\sigma_{\vec{p}}^{\vec{B}}$ is the substitution given by $\sigma(p_j) = B_j$ for $1 \leq j \leq n$ and $\sigma(x) = x$ for $x \notin \vec{p}$. A formula A is *classically right normal* for $\mathcal{R}$ if for every $k \geq 0$ there are formulae $B_1, \ldots, B_n$ such that $\vdash_{\mathsf{GcCutCon}\mathcal{R}} \bigvee_{i=1}^k A^i \leftrightarrow A\sigma_{\vec{p}}^{\vec{B}}$ with A^i and $\sigma_{\vec{p}}^{\vec{B}}$ as above (or equivalently if $A \to \bot$ is left normal). It is *intuitionistically right normal* for $\mathcal{R}$ if the above holds for $k \in \{0, 1\}$ and Gi instead of Gc. A context restriction $\langle F_1, F_2\rangle$ is *(intuitionistically or classically) normal* for $\mathcal{R}$ if every formula in F_1 (resp. F_2) is (intuitionistically or classically) left (resp. right) normal for $\mathcal{R}$.

Example 3.2.20. The formula p is intuitionistically and classically left and right normal for the set $\mathsf{G[ci]}$ of propositional rules and thus for every extension of $\mathsf{G[ci]}$. The formula $\Box p$ is intuitionistically and classically left normal for the rule set $\mathcal{R}_{\mathsf{K}}$ (and thus for all its extensions), since $\vdash_{\mathsf{G[ci]}\mathcal{R}_{\mathsf{K}}\mathsf{CutCon}} \bigwedge_{i=1}^{k} \Box p_i \leftrightarrow \Box \bigwedge_{i=1}^{k} p_i$ for every $k \geq 0$. It is not right normal for $\mathcal{R}_{\mathsf{K}}$. In the classical case furthermore the formula $\neg\Box p$ is right normal for $\mathcal{R}_{\mathsf{K}}$.

If we have a rule where all the restrictions are normal and contain only finitely many formulae, then intuitively we can absorb multiple instances of the same context formula into one proto rule. This is made precise in the following definition and lemma.

Definition 3.2.21. A context restriction $\langle F_1, F_2\rangle$ is *finite* if both sets F_1 and F_2 are finite. A *rule with finite context restrictions* is a rule in which every context restriction is finite.

Lemma and Definition 3.2.22 (ci)**.** *If $\mathcal{R}$ is a set of rules with context restrictions and R is a rule whose context restrictions $\{\mathcal{C}_1, \dots, \mathcal{C}_n\}$ are all finite and normal for $\mathcal{R}$, then the set of proto rules for R is equivalent over* $\mathsf{G[ci]}\mathcal{R}\mathsf{CutCon}$ *to the* canonical proto rule $\widehat{R}$ for R *given by the sequent $\Gamma_1 \Rightarrow \Gamma_2$ where for $i \in \{1, 2\}$ we set $\Gamma_i := \{C \mid C \in F_i$ where $\langle F_1, F_2\rangle = \mathcal{C}_j$ for some $j \leq n\}$. Thus R itself is equivalent over $\mathcal{R}$ to $\widehat{R}$.*

Proof. The proof of the first claim is straightforward using the definition of a normal restriction and the rules in $\mathsf{G[ci]}\mathcal{R}\mathsf{CutCon}$. The second statement follows easily from the definitions and the first claim. ☐

Example 3.2.23.

1. Since the only restriction occurring in the rule K_n is $\mathcal{C}_\emptyset$, the canonical proto rule for K_n is given by the empty context $\Rightarrow$.
2. The canonical proto rule for the rule R_{T} is given by the context $r \Rightarrow s$ and has the form
$$\frac{r, p \Rightarrow s}{r, \Box p \Rightarrow s}\ .$$
3. The canonical proto rule for the rule R_4 is given by the context $\Box r \Rightarrow$ and has the form
$$\frac{\Box r \Rightarrow p}{\Box r \Rightarrow \Box p}\ .$$

Since the context formulae for a translatable clause directly correspond to the context formulae of the corresponding proto rule, this gives us sufficient criteria for translatability of an axiom into a rule with restrictions.

Theorem 3.2.24 (ci)**.** *Let $\mathcal{R}$ be a rule set and let A be a (intuitionistically or classically) translatable clause with context formulae in (C_ℓ, C_r). If every formula in C_ℓ is left normal for*

$\mathcal{R}$ and every formula in C_r is right normal for $\mathcal{R}$, then A is equivalent over $\mathcal{R}$ to a rule with (finite) context restrictions.

Proof. Immediate from Theorem 3.2.14 and the preceding Lemma using the fact that the formulae on the left (resp. right) hand side of the context of the proto rule from the translation are in C_ℓ (resp. C_r). $\square$

Example 3.2.25.
1. The context formula $\Box q$ is (intuitionistically and classically) left normal for $\mathcal{R}_{\mathsf{K}}$ and thus translating the axiom (4) $\Box q \to \Box\Box q$ using Theorem 3.2.24 yields the well-known rule $R_{\mathsf{4}} = \{(\Rightarrow p; \langle\{\Box p\}, \emptyset\rangle)\}/ \Rightarrow \Box p$. This rule is thus equivalent to the axiom (4) over $\mathcal{R}_{\mathsf{K}}$.
2. Similarly, since variables are both left and right normal for $\mathcal{R}_{\mathsf{K}}$ translating the axiom (T) yields the standard rule $R_{\mathsf{T}} = \{(p \Rightarrow ; \mathcal{C}_{\mathsf{id}})\}/\Box p \Rightarrow$, which is equivalent to (T) over $\mathcal{R}_{\mathsf{K}}$.

In view of the fact that propositional variables are both left and right normal for every rule set including the propositional rules this immediately yields translation results for non-iterative and rank-1 axioms.

Corollary 3.2.26 (ci)**.** *In the intuitionistic case every translatable clause with context formulae in $(\emptyset, \emptyset)$ is equivalent over* Mon *to a shallow rule. If the clause is rank-1, then it is equivalent to a one-step rule. In the classical case every non-iterative (resp. rank-1) clause is equivalent over* Mon *to a shallow (resp. one-step) rule, and thus every non-iterative (resp. rank-1) formula is equivalent to a finite set of shallow (resp. one-step) rules.* $\square$

Remark 3.2.27. Since in the case that there are no context formulae we do not need to control which sides of the premisses the context formulae end up, the classical part of Corollary 3.2.26 stating that every non-iterative (resp. rank-1) formula whatsoever is translatable into an equivalent finite set of shallow (resp. one-step) rules holds for non-monotone modalities as well. In the intuitionistic case we still need to make sure that all the propositional connectives are invertible.

While Theorem 3.2.24 gives us sufficient criteria for when a clause is translatable into a single rule, the condition that all context formulae be normal is relatively strict. In particular it precludes the treatment of examples such as the (classical) axiom (5$\Box$) $\neg\Box p \to \Box\neg\Box p$. Moreover, if we only allow normal context formulae in the clauses we cannot construct rules with non-normal context restrictions by translation. On the other hand this can be done if instead of single translatable clauses we consider sets of formulae which are generated by a single translatable clause in a particular way.

Definition 3.2.28 (ci)**.** Let A be a classically translatable clause with context formulae in (C_ℓ, C_r) for $C_\ell = \{C_1, \dots, C_n\}$ and $C_r = \{D_1, \dots, D_m\}$ and variables in V. The *classical ω-set for A* is the set

$$\{A_{s_1,\dots,s_n,t_1,\dots,t_m} \mid s_i \geq 0, t_j \geq 0 \text{ for } i \leq n, j \leq m\} ,$$

where the formulae $A_{s_1,\dots,s_n,t_1,\dots,t_m}$ are constructed from A as follows: for $k \leq n$ every occurrence of the context formula C_k in A is replaced by the formula $\bigwedge_{i=1}^{s_k} C_k^i$, where again C_k^i is constructed from C_k by renaming the variables $\vec{p}$ of C_k to fresh variables $\vec{p}^i$; and for $k \leq m$ every occurrence of the context formula D_k is replaced by the formula $\bigvee_{i=1}^{t_k} D_k^i$, where again D_k^i is the result of substituting fresh variables for the variables in D_k. If A is an intuitionistically translatable clause, then the *intuitionistic ω-set for A* is the set

$$\{A_{s_1,\dots,s_n,t_1,\dots,t_m} \mid s_i \geq 0, t_j \in \{0,1\} \text{ for } i \leq n, j \leq m\} ,$$

where the formulae $A_{s_1,\dots,s_n,t_1,\dots,t_m}$ are constructed as in the classical case.

Note that in the intuitionistic case a translatable clause has at most one context formula in C_r, and thus at most one formula different from $\bot$ occurs in the disjunction on the right hand side of the top-level implication in the formulae $A_{s_1,\dots,s_n,t_1,\dots,t_m}$. Moreover, since no two formulae in $C_\ell \cup C_r$ share any variables with each other or with V, the ω-set for a translatable clause is in fact unique. Thus once the context formulae are specified it makes sense to speak of *the* ω-set for a translatable clause.

Example 3.2.29.
1. The (classical or intuitionistic) ω-set for the clause $(\mathsf{4})$ $\Box s \to \Box\Box s$ with context formulae in $(\{\Box s\}, \emptyset)$ is the set $\{\bigwedge_{i=1}^n \Box s^i \to \Box \bigwedge_{i=1}^n \Box s^i \mid n \geq 0\}$.
2. The classical ω-set for the clause $(\mathsf{5}\Box)$ $\neg\Box q \to \Box\neg\Box q$ with context formulae in $(\emptyset, \{\Box q\})$ is the set $\{\neg \bigvee_{i=1}^m \Box q^i \to \Box\neg \bigvee_{i=1}^m \Box q^i \mid m \geq 0\}$. If the context formulae are taken to be $(\{\neg\Box q\}, \emptyset)$, then the ω-set is $\{\bigwedge_{i=1}^n \neg\Box q^i \to \Box \bigwedge_{i=1}^n \neg\Box q^i \mid n \geq 0\}$.

Intuitively every conjunction resp. disjunction for a context formula in a formula of an ω-set corresponds to a fixed number of context formulae in a proto rule, and all the proto rules corresponding to formulae in the ω-set together are equivalent to a rule with restrictions.

Theorem 3.2.30 (ci)**.** *Every ω-set for a translatable clause is equivalent over* Mon *to a rule with (finite) context restrictions.*

Proof. By Theorem 3.2.24 every formula $A_{s_1,\dots,s_n,t_1,\dots,t_m}$ in the ω-set for a translatable clause A with context formulae in $(\{C_1,\dots,C_n\},\{D_1,\dots,D_m\})$ and variables in V is equivalent to a proto rule given by the context $\bigwedge_{i=1}^{s_1} C_1^i, \dots, \bigwedge_{i=1}^{s_n} C_n^i \Rightarrow \bigvee_{i=1}^{t_1} D_1^i, \dots, \bigvee_{i=1}^{t_m} D_m^i$. Using the rules in $\mathsf{G[ci]CutCon}$ this is equivalent to the proto rule given by the context with $C_k^1,\dots,C_k^{s_k}$ instead of $\bigwedge_{i=1}^{s_k} C_k^i$ resp. $D_k^1,\dots,D_k^{t_k}$ instead of $\bigvee_{i=1}^{t_k} D_k^i$. Note that for the latter case in the intuitionistic setting we only need to deal with at most one formula D_k^j, where we use the fact that empty disjunctions are defined as $\bot$. Since all the proto rules come from axioms which are generated by the same translatable clause they are proto rules for the same rule with restrictions. Now the result follows immediately. $\Box$

The technique of translating ω-sets of axioms instead of single axioms now allows us to treat examples like the axiom (5) as well.

Example 3.2.31.

1. Translating the ω-set $\{\neg \bigvee_{i=1}^{n} \Box q^i \to \Box\neg \bigvee_{i=1}^{n} \Box q^i \mid n \geq 0\}$ for the axiom (5$\Box$) $\neg\Box q \to \Box\neg\Box q$ using Theorem 3.2.30 yields first the set of proto rules for the rule $R = \{(\Rightarrow p; \langle\emptyset, \{\Box p\}\rangle)\}/ \Rightarrow \Box p$ given by the sequents $\Rightarrow \Box q^1, \dots, \Box q^n$ for $n \geq 0$. Together, all these proto rules are are equivalent to the rule R itself.

2. By classical propositional reasoning and the axioms of K adding both axioms (4) $\Box q_\ell \to \Box\Box q_\ell$ and (5$\Box$) $\neg\Box q_r \to \Box\neg\Box q_r$ is equivalent to adding the set

$$\left\{ \bigwedge_{i=1}^{n} \Box q_\ell^i \wedge \neg \bigvee_{j=1}^{m} \Box q_r^j \to \Box(\bigwedge_{i=1}^{n} \Box q_\ell^i \wedge \neg \bigvee_{j=1}^{m} \Box q_r^j) \mid m, n \geq 0 \right\}$$

   of axioms, which is an ω-set for the classically translatable clause $\Box q_\ell \wedge \neg\Box q_r \to \Box(\Box q_\ell \wedge \neg\Box q_r)$. Using the method of Theorem 3.2.30 this set translates into the standard rule $R_{45} = \{(\Rightarrow p; \mathcal{C}_{45})\}/ \Rightarrow \Box p$ with restriction $\mathcal{C}_{45} = \langle\{\Box p\}, \{\Box p\}\rangle$ from Table 2.4. Thus the two axioms (4) and (5$\Box$) together are equivalent over $\mathcal{R}_{\mathsf{K}}$ to the rule R_{45}.

Corollary 3.2.32 (ci)**.** *Let $\mathcal{R}$ be a set of rules with restrictions and let A be an axiom. If A is equivalent over $\mathcal{R}$ to an ω-set for a translatable clause, then A is equivalent over $\mathcal{R}$ to a rule with (finite) restrictions.* $\Box$

Remark 3.2.33. We may also use the previous corollary to give a slightly different proof of Theorem 3.2.24: Whenever the context formulae for a translatable clause are normal, then the clause is equivalent to its ω-set. Now the corollary yields equivalence of the clause to a rule.

3.3 From Rules to Axioms

The results of the previous section raise the question whether the format of ω-sets for axioms is really necessary. It turns out that for monotone modalities the format is both necessary and sufficient in the sense that an axiom can be translated into a rule with finite restrictions if and only if adding the axiom is equivalent to adding an ω-set. The restriction to rules with finite context restrictions here is necessary if we aim for a single ω-set. On the other hand, using the following lemma it is clear that an axiom is equivalent to a set of rules with restrictions if and only if it is axiomatically equivalent to the union of a set of ω-sets.

Lemma 3.3.1 (ci)**.** *Every rule with restrictions is equivalent to a set of rules with finite restrictions.*

Proof. Let R be a rule with restrictions $\{\langle F_1^i, F_2^i\rangle \mid 1 \le i \le n\}$. The rules in the rule set $\mathcal{R}_R$ are then constructed by replacing for each restriction the components F_1^i and F_2^i by finite subsets $G_1^i \subseteq F_1^i$ and $G_2^i \subseteq F_2^i$. Then for every application of the rule R there is a rule in $\mathcal{R}_R$ whose context restrictions include all the relevant formulae, and for the other direction every application of a rule in $\mathcal{R}_R$ is simulated by Weakening of some of the premisses and an application of the rule R. □

Thus in the following we will concentrate on rules with finite restrictions. We show the correspondence to ω-sets for such rules by translating them back into ω-sets of axioms. The main idea for this is to turn the premisses and the conclusion of a proto rule into formulae, and then construct the axiom out of the formula corresponding to the conclusion by suitably incorporating the premisses. The latter step is accomplished following [Sch07] using a suitably adjusted notion of a projective formula [Ghi99] and making use of a carefully chosen substitution witnessing the projectivity of the formula corresponding to the premisses of the proto rule. For the purpose of constructing this substitution we would like the premisses of the original rule to be of a specific form.

Definition 3.3.2 (ci)**.** A rule $\mathcal{P}/\Sigma \Rightarrow \Pi$ is in *standard form* if

1. no variable occurs both on the left hand side of a premiss and on the right hand side of a (possibly different) premiss
2. whenever $(\Rightarrow p; \mathcal{C}_\emptyset) \in \mathcal{P}$ then there is no premiss $(\Gamma \Rightarrow p, \Delta; \mathcal{C}) \in \mathcal{P}$ with $\mathcal{C} \neq \mathcal{C}_\emptyset$ or $\Gamma \cup \Delta \neq \emptyset$
3. there is no premiss $(\Rightarrow ; \mathcal{C}) \in \mathcal{P}$ for any restriction $\mathcal{C}$
4. in the intuitionistic case, there is a formula D such that the right component of every restriction occurring in $\mathcal{P}$ is $\emptyset$ or $\{D\}$.

Fortunately, for monotone modalities every rule with finite restrictions can be manipulated in such a way that it becomes a rule in standard form (or a finite set of such rules in the intuitionistic case). This is shown in the following lemmata. We first consider Property 4 for the intuitionistic setting.

Lemma 3.3.3 (i)**.** *Every rule with finite restrictions is equivalent over* Mon *to a finite set of rules with restrictions satisfying Property 4 of Definition 3.3.2 above.*

Proof. Suppose $R = \mathcal{P}/\Sigma \Rightarrow \pi$. If $\pi \neq \emptyset$, then no application of R has context formulae on the right hand side of its conclusion, and w.l.o.g. the right component of every restriction occurring in $\mathcal{P}$ is empty. So suppose that $\pi = \emptyset$ and we have premisses $\mathcal{P} = \{\Gamma_i \Rightarrow \delta_i; \langle F_i, G_i\rangle \mid 1 \le i \le n\}$. Since the G_i are finite, the set $\bigcup_{i=1}^n G_i$ is finite as well. W.l.o.g. we assume that for all $A, B \in \bigcup_{i=1}^n G_i$ with $A \neq B$ we have $\mathsf{var}(A) \cap \mathsf{var}(B) = \emptyset$. This can always be achieved using injective renamings. We would like to construct formulae $C_1, \dots, C_m$ from the formulae in $\bigcup_{i=1}^n G_i$ such that we can replace the rule R by m rules $R_1, \dots, R_m$ where the right components of restrictions of rule R_j are either $\emptyset$ or $\{C_j\}$. The idea is to construct all possible "minimal" formulae satisfying a given subset of the restrictions. For this we make use of (syntactic) *unification* [EFT94, EFT96]. Remember that a *unifier* for a set $\{B_1, \dots, B_k\}$ of formulae is a substitution σ such that $B_1\sigma = \dots = B_k\sigma$. If a unifier for a set of formulae exists, then we call the set *unifiable.* A *most general unifier* for a set S of formulae is a unifier mgu_S such that for every other unifier σ of S there is a substitution τ with $\sigma = \tau \circ \mathsf{mgu}_S$. If such a most general unifier exists, then in fact it is unique, so we also speak of *the* most general unifier. A straightforward application of the standard algorithm for syntactic unification as given e.g. in [EFT94, EFT96] shows that whenever a set of formulae is unifiable, then there is a most general unifier. Now for every non-empty unifiable set $S \subseteq \bigcup_{i=1}^n G_i$ define $C_S := B\,\mathsf{mgu}_S$ where $B \in S$ arbitrary. Furthermore for such an S define the premisses $\mathcal{P}_S$ by

$$\mathcal{P}_S := \begin{array}{l}\{(\Gamma_i \Rightarrow \delta_i; \langle F_i, \{C_S\}\rangle) \mid 1 \leq i \leq n,\ S \cap G_i \neq \emptyset\} \\ \cup \{(\Gamma_i \Rightarrow \delta_i; \langle F_i, \emptyset\rangle) \mid 1 \leq i \leq n,\ S \cap G_i = \emptyset\}\end{array}$$

and define the rule R_S by $R_S := \mathcal{P}_S/\Sigma \Rightarrow \pi$. We claim that then the original rule R is equivalent to the set $\{R_S \mid S \subseteq \bigcup_{i=1}^n G_i,\ S \text{ unifiable}\}$. To see why this is the case first consider an application of the rule R with context $\Theta \Rightarrow A$. If A is not a substitution instance of any formula in $\bigcup_{i=1}^n G_i$, then we may replace the application by an application of the rule R_S for an arbitrary unifiable $S \subseteq \bigcup_{i=1}^n G_i$. Otherwise let $I \subseteq \{1, \dots, n\}$ be the set of indices i such that A is a substitution instance of a formula in G_i and for each $i \in I$ choose one such formula $B_i \in G_i$. Then for every $i \in I$ the formula A is also a substitution instance of the formula $B_i\,\mathsf{mgu}_{\{B_i \mid i \in I\}}$ and thus the application of R can be replaced by an application of the rule $R_{\{B_i \mid i \in I\}}$. For the other direction suppose that $S \subseteq \bigcup_{i=1}^n G_i$ is unifiable and consider an application of the rule R_S given by the substitution σ and the context $\Theta \Rightarrow A$. If A is a substitution instance of the formula C_S, then by definition of C_S for every G_i with $G_i \cap S \neq \emptyset$ there is a formula $B \in G_i$ such that A is a substitution instance of B. Moreover, for every premiss $(\Gamma_j \Rightarrow\ ; \langle F_j, G_j\rangle)$ with $G_j \cap S = \emptyset$ we have a derivation of $\Gamma_j\sigma, \Theta \upharpoonright_{F_j} \Rightarrow$ from which using W we can easily derive the sequent $\Gamma_j\sigma, \Theta \upharpoonright_{F_j} \Rightarrow A \upharpoonright_{G_j}$. Thus we may replace the application of R_S by an application of R. Finally, the set $\{R_S \mid S \subseteq \bigcup_{i=1}^n G_i,\ S \text{ unifiable}\}$ is a finite set of rules, since all restrictions of R were finite. $\square$

Lemma 3.3.4 (ci). *Every non-trivial rule with finite restrictions is equivalent over* Mon *to a finite set of rules with finite restrictions in standard form. If the rule was shallow (resp. one-step), then so are the equivalent rules in standard form.*

Proof. We show how to equivalently transform an arbitrary rule into a rule satisfying properties 1-4 in Definition 3.3.2 characterising the standard form. Suppose we have a rule $R = \mathcal{P}/\Sigma \Rightarrow \Pi$.

If we are in the intuitionistic setting, we first take care of property 4 using Lemma 3.3.3 and then consider each rule from the resulting finite set separately. The remaining procedure is the same in both the intuitionistic and the classical case. The next step is to make the rule satisfy the first property. This is done by using the fact that all our rule sets include the rules Mon and replacing the rule by the cut between this rule and all possible rules $\mathsf{Mon}_\heartsuit$ according to Definition 2.4.1. In the resulting rule no variable occurs in the premisses both on the left hand side of a sequent and on the right hand side of a sequent, and Lemma 2.4.5 ensures that this rule is equivalent over Mon to the original rule. Note that in the intuitionistic case this does not destroy property 4 since all non-empty right components of context restrictions occurring in the premisses are the same. For the second property suppose that there are premisses $(\Rightarrow p; \mathcal{C}_\emptyset)$ and $(\Gamma \Rightarrow p, \Delta; \mathcal{C})$ in $\mathcal{P}$, where $\Gamma \cup \Delta \neq \emptyset$ or $\mathcal{C} \neq \mathcal{C}_\emptyset$. Then instances of the latter are derived by Weakening from instances of $(\Rightarrow p; \mathcal{C}_\emptyset)$ and thus the premiss $(\Gamma \Rightarrow p, \Delta; \mathcal{C})$ can be omitted. Finally, if we have a premiss $(\Rightarrow ; \mathcal{C})$, then the rule is subsumed by Weakening and therefore trivial. □

For the rest of this section we assume w.l.o.g. that for monotone modalities all rules with finite restrictions are in standard form. While our goal is to translate such rules into axioms, the varying number of context formulae makes it hard to translate rules directly. For this reason again we first consider proto rules. Given a proto rule the first step is to turn its premisses and conclusion into formulae.

Definition 3.3.5 (ci)**.** Let $R = \mathcal{P}/\Sigma \Rightarrow \Pi$ be a rule and $\widehat{R}$ a proto rule for R given by the context $\Gamma \Rightarrow \Delta$. The formulae $\mathsf{Prem}_{\widehat{R}}$ and $\mathsf{Concl}_{\widehat{R}}$ are defined by

$$\begin{aligned}\mathsf{Prem}_{\widehat{R}} &:= \bigwedge_{(\Theta \Rightarrow \Xi; \langle F_1; F_2 \rangle) \in \mathcal{P}} \Big(\bigwedge \Gamma \restriction_{F_1} \wedge \bigwedge \Theta \to \bigvee \Delta \restriction_{F_2} \vee \bigvee \Xi \Big) \\ \mathsf{Concl}_{\widehat{R}} &:= \bigwedge \Gamma \wedge \bigwedge \Sigma \to \bigvee \Delta \vee \bigvee \Pi \,.\end{aligned}$$

where as usual we take empty conjunctions to be $\top$ and empty disjunctions to be $\bot$. Note that in the intuitionistic case since the sequents are asymmetric the disjunctions in the succedent of the implications contain at most one formula other than $\bot$.

Then by propositional reasoning it is clear that the premisses of a proto rule $\widehat{R}$ (resp. its conclusion) are derivable if and only if the sequent $\Rightarrow \mathsf{Prem}_{\widehat{R}}$ (resp. $\Rightarrow \mathsf{Concl}_{\widehat{R}}$) is derivable.

Example 3.3.6. 1. The proto rule for the rule R_4 given by the context $\Box p_1, \Box p_2 \Rightarrow$ is

$$\widehat{R_4} = \frac{\Box p_1, \Box p_2 \Rightarrow p}{\Box p_1, \Box p_2 \Rightarrow \Box p}$$

and gives $\mathsf{Prem}_{\widehat{R_4}} = \Box p_1 \wedge \Box p_2 \wedge \top \to \bot \vee p$ and $\mathsf{Concl}_{\widehat{R_4}} = \Box p_1 \wedge \Box p_2 \wedge \top \to \bot \vee \Box p$.

2. The proto rule for the rule R_{T} given by the context $p_1, p_2 \Rightarrow$ is

$$\widehat{R_{\mathsf{T}}} = \frac{p_1, p_2, p \Rightarrow}{p_1, p_2, \Box p \Rightarrow}$$

and gives the formulae $\mathsf{Prem}_{\widehat{R_{\mathsf{T}}}} = p_1 \wedge p_2 \wedge p \to \bot \vee \bot$ and $\mathsf{Concl}_{\widehat{R_{\mathsf{T}}}} = p_1 \wedge p_2 \wedge \Box p \to \bot \vee \bot$.

3. The proto rule for the rule R_{45} given by the context $\Box p_1 \Rightarrow \Box q_1$ is

$$\widehat{R_{\mathsf{45}}} = \frac{\Box p_1 \Rightarrow p, \Box q_1}{\Box p_1 \Rightarrow \Box p, \Box q_1}$$

and gives $\mathsf{Prem}_{\widehat{R_{\mathsf{45}}}} = \Box p_1 \wedge \top \to \Box q_1 \vee p$ and $\mathsf{Concl}_{\widehat{R_{\mathsf{45}}}} = \Box p_1 \wedge \top \to \Box q_1 \vee \Box p$.

For capturing the information given in the premisses and injecting it into the formula $\mathsf{Concl}_{\widehat{R}}$ we follow [Sch07] and make use of slightly adapted notions from the theory of projective formulae, see e.g. [Ghi99].

Definition 3.3.7 (ci)**.** A formula $A \in \mathcal{F}(\Lambda)$ is *projective* if there is a substitution $\sigma : \mathsf{Var} \to \mathcal{F}(\Lambda)$ such that $\vdash_{\mathsf{G[ci]MonCutConW}} \Rightarrow A\sigma$ and such that for all $p \in \mathsf{var}(A)$ we have $\vdash_{\mathsf{G[ci]MonCutConW}} A \Rightarrow p \leftrightarrow p\sigma$. Such a substitution *witnesses projectivity of* A.

It is now standard to show the following lemma.

Lemma 3.3.8 (ci)**.** *If $A \in \mathcal{F}(\Lambda)$ is a formula and $\sigma : \mathsf{Var} \to \mathcal{F}(\Lambda)$ is a substitution witnessing projectivity of A, then for every formula B with $\mathsf{var}(B) \subseteq \mathsf{var}(A)$ we have $\vdash_{\mathsf{G[ci]MonCutConW}} A \Rightarrow B \leftrightarrow B\sigma$.*

Proof. By induction on the complexity of the formula B using the monotonicity rules. $\Box$

Given a proto rule $\widehat{R}$ once we have a substitution witnessing the projectivity of the formula $\mathsf{Prem}_{\widehat{R}}$ we are done using the following lemma.

Lemma 3.3.9 (ci)**.** *If $\widehat{R}$ is a proto rule and σ a substitution witnessing projectivity of $\mathsf{Prem}_{\widehat{R}}$, then the axiom $\mathsf{Concl}_{\widehat{R}}\sigma$ is equivalent to $\widehat{R}$ over $\mathsf{Mon}\mathcal{R}$ for every rule set $\mathcal{R}$.*

Proof. Let $\widehat{R}$ be a proto rule with premisses $\Gamma_1 \Rightarrow \Delta_1, \dots, \Gamma_n \Rightarrow \Delta_n$ and conclusion $\Sigma \Rightarrow \Pi$, and let σ be a substitution witnessing projectivity of $\mathsf{Prem}_{\widehat{R}}$. We first show that the proto rule $\widehat{R}$ is a derivable rule using the ground sequent $\Rightarrow \mathsf{Concl}_{\widehat{R}}\sigma$. By the fact that σ witnesses projectivity of $\mathsf{Prem}_{\widehat{R}}$ we know that there is a derivation $\mathcal{D}$ in $\mathsf{G[ci]MonCutConW}\mathcal{R}$ of the sequent $\mathsf{Prem}_{\widehat{R}} \Rightarrow \mathsf{Concl}_{\widehat{R}}\sigma \to \mathsf{Concl}_{\widehat{R}}$, and furthermore the sequent $\Rightarrow \mathsf{Concl}_{\widehat{R}}\sigma$ is a ground sequent and thus derivable. But then we may replace the proto rule $\widehat{R}$ by the following derivation:

$$\dfrac{\dfrac{}{\Rightarrow \mathsf{Concl}_{\widehat{R}}\sigma} \qquad \dfrac{\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1 \quad \dots \quad \Gamma_n \Rightarrow \Delta_n}{\Rightarrow \mathsf{Prem}_{\widehat{R}}}\,\mathsf{G[ci]W} \qquad \begin{array}{c}\mathcal{D}\\ \vdots \\ \mathsf{Prem}_{\widehat{R}} \Rightarrow \mathsf{Concl}_{\widehat{R}}\sigma \to \mathsf{Concl}_{\widehat{R}}\end{array}}{\dfrac{\Rightarrow \mathsf{Concl}_{\widehat{R}}\sigma \to \mathsf{Concl}_{\widehat{R}}}{\mathsf{Concl}_{\widehat{R}}\sigma \Rightarrow \mathsf{Concl}_{\widehat{R}}}\,\mathsf{G[ci]Cut}}\,\mathsf{Cut}}{\dfrac{\Rightarrow \mathsf{Concl}_{\widehat{R}}}{\Sigma \Rightarrow \Pi}\,\mathsf{G[ci]Cut}}\,\mathsf{Cut}$$

This derivation is easily modified to accommodate the additional substitution and context in an application of the proto rule $\widehat{R}$.

For the other direction we need to derive the sequent $\mathsf{Concl}_{\widehat{R}}\sigma$ using $\widehat{R}$. By projectivity we have $\vdash_{\mathsf{G[ci]MonCutCon}\mathcal{R}} \Rightarrow \mathsf{Prem}_{\widehat{R}}\sigma$. Now resolving the propositional connectives using $\mathsf{G[ci]Cut}$ yields the sequents $\Gamma_i\sigma \Rightarrow \Delta_i\sigma$, and applying the proto rule $\widehat{R}$ and propositional rules gives $\vdash_{\mathsf{G[ci]MonCutConW}\mathcal{R}} \Rightarrow \mathsf{Concl}_{\widehat{R}}\sigma$. ☐

In particular in the intuitionistic case it is not entirely clear that we can always construct such a substitution. Fortunately, the premisses of rules in standard form have a distinct shape which ensures that this is possible. We first consider the (easier) classical case.

Definition 3.3.10 (c)**.** Let $R = \mathcal{P}/\Sigma \Rightarrow \Pi$ be a rule in standard form and $\widehat{R}$ a proto rule for R given by $\Gamma \Rightarrow \Delta$. Define the substitution $\theta^{\mathsf{c}}_{\widehat{R}}$ by

$$\theta^{\mathsf{c}}_{\widehat{R}}(p) := \begin{cases} \top & : (\Rightarrow p; \mathcal{C}_\emptyset) \in \mathcal{P} \\ \mathsf{Prem}_{\widehat{R}} \to p & : (\Theta \Rightarrow p, \Xi; \mathcal{C}) \in \mathcal{P} \text{ for some } \Theta \Rightarrow \Xi \text{ and } \mathcal{C} \text{ with } \Theta \cup \Xi \neq \emptyset \text{ or } \mathcal{C} \neq \mathcal{C}_\emptyset \\ \mathsf{Prem}_{\widehat{R}} \wedge p & : (\Theta, p \Rightarrow \Xi; \mathcal{C}) \in \mathcal{P} \text{ for some } \Theta \Rightarrow \Xi \\ p & : \text{otherwise.} \end{cases}$$

The substitution $\theta_{\widehat{R}}$ is well-defined, since in a rule in standard form no variable occurs both on the left side of a premiss and on the right side of a premiss, and if it occurs on the right side, then it occurs either only in a premiss of the form given in the first case or only in premisses of the form given in the second case.

Lemma 3.3.11 (c)**.** *If $\widehat{R}$ is a proto rule for a rule R in standard form, then the substitution $\theta^{\mathsf{c}}_{\widehat{R}}$ witnesses projectivity of* $\mathsf{Prem}_{\widehat{R}}$.

Proof. Suppose that the proto rule $\widehat{R}$ for the rule $R = \mathcal{P}/\Sigma \Rightarrow \Pi$ is given by the context $\Gamma \Rightarrow \Delta$. It is easy to see that for every variable $p \in \mathsf{var}\left(\mathsf{Prem}_{\widehat{R}}\right)$ we have $\vdash_{\mathsf{GcMonCutConW}} \mathsf{Prem}_{\widehat{R}} \Rightarrow p \leftrightarrow p\theta^{\mathsf{c}}_{\widehat{R}}$. To see that $\vdash_{\mathsf{GcMonCutConW}} \Rightarrow \mathsf{Prem}_{\widehat{R}}\theta^{\mathsf{c}}_{\widehat{R}}$ consider a premiss $(\Theta \Rightarrow \Xi; \langle F_1, F_2\rangle)$ from $\mathcal{P}$ and the corresponding clause $(\bigwedge \Gamma \restriction_{F_1} \wedge \bigwedge \Theta \to \bigvee \Delta \restriction_{F_2} \vee \bigvee \Xi)\theta^{\mathsf{c}}_{\widehat{R}}$ from $\mathsf{Prem}_{\widehat{R}}\theta^{\mathsf{c}}_{\widehat{R}}$. If the premiss has the form $(\Rightarrow p; \mathcal{C}_\emptyset)$, then this clause only consists of the formula $\top$ which is trivially derivable. Otherwise, since the substitution $\theta^{\mathsf{c}}_{\widehat{R}}$ is the identity on context formulae this clause is the same as $\bigwedge \Gamma \upharpoonright_{F_1} \wedge \bigwedge \Theta\theta^{\mathsf{c}}_{\widehat{R}} \to \bigvee \Delta \upharpoonright_{F_2} \vee \bigvee \Xi\theta^{\mathsf{c}}_{\widehat{R}}$. Now if $\Theta \neq \emptyset$, then this is equivalent to $\bigwedge \Gamma \upharpoonright_{F_1} \wedge \mathsf{Prem}_{\widehat{R}} \wedge \Theta \to \bigvee \Delta \upharpoonright_{F_2} \vee A$, where A is either $\bot$ if Ξ is empty, or equivalent to $\bigvee \Xi$ or $\mathsf{Prem}_{\widehat{R}} \to \bigvee \Xi$ if $\Xi \neq \emptyset$. Since $\bigwedge \Gamma \upharpoonright_{F_1} \wedge \bigwedge \Theta \to \bigvee \Delta \upharpoonright_{F_2} \vee \bigvee \Xi$ is a clause in $\mathsf{Prem}_{\widehat{R}}$ it is easy to see that in both cases the clause is derivable. On the other hand, if $\Theta = \emptyset$, then the clause is equivalent to $\bigwedge \Gamma \upharpoonright_{F_1} \to \bigvee \Delta \upharpoonright_{F_2} \vee (\mathsf{Prem}_{\widehat{R}} \to \bigvee \Xi)$, which again is easily seen to be derivable using the fact that $\mathsf{Prem}_{\widehat{R}}$ contains the clause $\bigwedge \Gamma \upharpoonright_{F_1} \to \bigvee \Delta \upharpoonright_{F_2} \vee \bigvee \Xi$. □

While the substitution defined above in principle also works in the intuitionistic case, we would like the formula $\mathsf{Concl}_{\widehat{R}}\theta_R$ to be a translatable clause in the sense of Definition 3.2.9 as well. But this means that in the intuitionistic setting we cannot substitute the formula $\mathsf{Prem}_{\widehat{R}} \to p$ for variables p occurring under a modality on the right hand side of the principal formulae, since it is not intuitionistically left resolvable. This problem can be fixed using the following substitution $\theta^{\mathsf{i}}_{\widehat{R}}$ instead.

Definition 3.3.12 (i)**.** For a rule $R = \mathcal{P}/\Sigma \Rightarrow \pi$ in standard form and a proto rule $\widehat{R}$ for R given by the context $\Gamma \Rightarrow \delta$ define the substitution $\theta^{i}_{\widehat{R}}$ by

$$\theta^{i}_{\widehat{R}}(p) := \begin{cases} \bigvee_{(\ \Rightarrow p; \langle F_1, F_2\rangle) \in \mathcal{P}} \bigwedge \Gamma \upharpoonright_{F_1} \vee p & : (\Theta \Rightarrow p; \langle F_1, F_2\rangle) \in \mathcal{P} \text{ for some } \Theta \text{ and } \langle F_1, F_2\rangle \neq \mathcal{C}_\emptyset \\ \mathsf{Prem}_{\widehat{R}} \wedge p & : (\Theta, p \Rightarrow \xi; \mathcal{C}) \in \mathcal{P} \text{ for some } \Theta \Rightarrow \xi \text{ and } \mathcal{C} \\ p & : \text{otherwise.} \end{cases}$$

Note that in the previous Definition if there is a premiss $(\ \Rightarrow p; \mathcal{C}_\emptyset)$ in $\mathcal{P}$, then the formula $\theta^{\mathsf{i}}_{\widehat{R}}(p)$ is equivalent to $\top$, and if there is a premiss $(\Theta \Rightarrow p; \mathcal{C})$ but no premiss $(\ \Rightarrow p; \mathcal{C}')$ in $\mathcal{P}$, then $\theta^{\mathsf{i}}_{\widehat{R}}(p)$ is equivalent to p.

Lemma 3.3.13 (i)**.** *If $\widehat{R}$ is a proto rule for a rule R in standard form, then the substitution $\theta^{\mathsf{i}}_{\widehat{R}}$ witnesses projectivity of* $\mathsf{Prem}_{\widehat{R}}$.

Proof. Similar to the classical case: Again, let the proto rule $\widehat{R}$ for the rule $R = \mathcal{P}/\Sigma \Rightarrow \pi$ be given by the context $\Gamma \Rightarrow \delta$. Again, standard intuitionistic propositional reasoning gives $\vdash_{\mathsf{GiMonCutConW}} \mathsf{Prem}_{\widehat{R}} \Rightarrow p \leftrightarrow p\theta^{\mathsf{i}}_{\widehat{R}}$. To show that $\vdash_{\mathsf{GiMonCutConW}} \mathsf{Prem}_{\widehat{R}}\theta^{\mathsf{i}}_{\widehat{R}}$ consider a premiss $(\Theta \Rightarrow \xi; \langle F_1, F_2\rangle)$ and the corresponding clause $\bigwedge \Gamma \upharpoonright_{F_1} \wedge \bigwedge \Theta\theta^{\mathsf{i}}_{\widehat{R}} \to \delta \upharpoonright_{F_2} \vee \xi\theta^{\mathsf{i}}_{\widehat{R}}$ from $\mathsf{Prem}_{\widehat{R}}\theta^{\mathsf{i}}_{\widehat{R}}$. The case for premisses $(\ \Rightarrow p; \mathcal{C}_\emptyset)$ is dealt with as in the classical case. Otherwise, if $\Theta \neq \emptyset$, this is equivalent to either $\bigwedge \Gamma \upharpoonright_{F_1} \wedge \mathsf{Prem}_{\widehat{R}} \wedge \bigwedge \Theta \to \delta \upharpoonright_{F_2} \vee \bot$ or $\bigwedge \Gamma \upharpoonright_{F_1} \wedge \mathsf{Prem}_{\widehat{R}} \wedge \bigwedge \Theta \to \bot \vee \bigvee_{(\ \Rightarrow \xi; \langle G_1, G_2\rangle) \in \mathcal{P}} \bigwedge \Gamma \upharpoonright_{G_1} \vee \xi$, depending on whether ξ or δ is empty. But both of these are intuitionistically derivable, since the clause $\bigwedge \Gamma \upharpoonright_{F_1} \wedge \bigwedge \Theta \to \delta \upharpoonright_{F_2} \vee \xi$ occurs in $\mathsf{Prem}_{\widehat{R}}$. Finally, if $\Theta = \emptyset$, then since the rule was in standard form δ must be empty and the clause is equivalent to $\bigwedge \Gamma \upharpoonright_{F_1} \to \bigvee_{(\ \Rightarrow \xi; \langle G_1, G_2\rangle) \in \mathcal{P}} \bigwedge \Gamma \upharpoonright_{G_1} \vee \xi$, which is easily seen to be derivable. □

Putting all the pieces together we obtain an automatic translation of proto rules into translatable clauses for monotone modalities.

Theorem 3.3.14 (ci)**.** *Every proto rule $\widehat{R}$ for a rule R in standard form given by a context $\Gamma \Rightarrow \Delta$ is equivalent over* Mon *to a translatable clause with context formulae in* $(\mathsf{Supp}(\Gamma), \mathsf{Supp}(\Delta))$*. The translation is automatic.*

Proof. Let R be a rule with restrictions in standard form and let $\widehat{R}$ be a proto rule for R given by the context $\Gamma \Rightarrow \Delta$ and let V be the set of variables occurring in the rule R. By Lemmata 3.3.9 and 3.3.11 resp. 3.3.13 the axiom $\mathsf{Concl}_{\widehat{R}}\theta_{\widehat{R}}^{[\mathsf{ci}]}$ is equivalent over Mon to $\widehat{R}$. From the definition of $\mathsf{Concl}_{\widehat{R}}\theta_{\widehat{R}}^{[\mathsf{ci}]}$ it is clear that the translation is automatic. To see that the formula $\mathsf{Concl}_{\widehat{R}}\theta_{\widehat{R}}^{[\mathsf{ci}]}$ is a translatable clause with context formulae in $(\mathsf{Supp}(\Gamma), \mathsf{Supp}(\Delta))$, for the sake of presentation assume that all modalities are unary. The general case is analogous. It is clear by construction that $\mathsf{Concl}_{\widehat{R}}$ is of the form $A_1 \wedge \ldots \wedge A_n \to B_1 \vee \cdots \vee B_m$, and that the formulae in Γ occur only on the top level in the A_i, those in Δ only in the B_i. Now consider a variable $p \in V$ occurring in a formula $A_i = \heartsuit p$ not in Γ. We need to check that the formula $\theta_{\widehat{R}}^{[\mathsf{ci}]}(p)$ is in $\mathcal{F}_r(\mathsf{Supp}(\Gamma), V, \mathsf{Supp}(\Delta)) \cup V$. If the variable p does not occur in the premisses, then we have $\theta_{\widehat{R}}^{[\mathsf{ci}]}(p) = p \in V$. Otherwise, since the rule was in standard form the variable p occurs only on the left hand side of the premisses, and thus we have $\theta_{\widehat{R}}^{[\mathsf{ci}]}(p) = \mathsf{Prem}_{\widehat{R}} \wedge p$. But by construction of $\mathsf{Prem}_{\widehat{R}}$ it is clear that the formula $\mathsf{Prem}_{\widehat{R}} \wedge p$ is in ${\mathcal{F}_r}^{[\mathsf{ci}]}(\mathsf{Supp}(\Gamma), V, \mathsf{Supp}(\Delta))$. Similarly, if the variable p occurs in a formula $B_i = \heartsuit p$ not in Δ, then if it does not occur in the premisses we have $\theta_{\widehat{R}}^{[\mathsf{ci}]}(p) = p \in V$. Otherwise it occurs only on the right hand side of the premisses. Thus we have $\theta_{\widehat{R}}^{\mathsf{c}}(p) = \mathsf{Prem}_{\widehat{R}} \to p$ and $\theta_{\widehat{R}}^{\mathsf{i}}(p) = \bigvee_{(\ \Rightarrow p;\mathcal{C}) \in \mathcal{P}} \bigwedge \Gamma \restriction_{\mathcal{C}} \vee p$. Again, both of these formulae are by construction in ${\mathcal{F}_\ell}^{\mathsf{c}}(\mathsf{Supp}(\Gamma), V, \mathsf{Supp}(\Delta))$ resp. ${\mathcal{F}_\ell}^{\mathsf{i}}(\mathsf{Supp}(\Gamma), V, \mathsf{Supp}(\Delta))$. Thus the formula $\mathsf{Concl}_{\widehat{R}}\theta_{\widehat{R}}^{[\mathsf{ci}]}$ is a translatable clause with context formulae in $(\mathsf{Supp}(\Gamma), \mathsf{Supp}(\Delta))$ and variables in V. $\square$

Example 3.3.15. Continuing Example 3.3.6 we have:

1. For the proto rule $\widehat{R_4}$ for the rule R_4 given by the context $\Box p_1, \Box p_2 \Rightarrow$ we get the substitution $\theta_{\widehat{R_4}}^{\mathsf{i}}$ with $\theta_{\widehat{R_4}}^{\mathsf{i}}(p) = \Box p_1 \wedge \Box p_2$ and $\theta_{\widehat{R_4}}^{\mathsf{i}}(p_i) = p_i$ for $i = 1, 2$. Thus the proto rule $\widehat{R_4}$ is equivalent over Mon to the axiom
$$\mathsf{Concl}_{\widehat{R_4}}\theta_{\widehat{R_4}}^{\mathsf{i}} = \Box p_1 \wedge \Box p_2 \wedge \top \to \bot \vee \Box(\Box p_1 \wedge \Box p_2)\ .$$

2. For the proto rule $\widehat{R_\mathsf{T}}$ for the rule R_T given by the context $p_1, p_2 \Rightarrow$ we get the substitution $\theta_{\widehat{R_\mathsf{T}}}^{[\mathsf{ci}]}$ with $\theta_{\widehat{R_\mathsf{T}}}^{[\mathsf{ci}]}(p) = (p_1 \wedge p_2 \wedge p \to \bot \vee \bot) \wedge p$ and $\theta_{\widehat{R_\mathsf{T}}}^{[\mathsf{ci}]}(p_i) = p_i$ for $i = 1, 2$. Thus the proto rule $\widehat{R_\mathsf{T}}$ is equivalent over Mon to the axiom
$$\mathsf{Concl}_{\widehat{R_\mathsf{T}}}\theta_{\widehat{R_\mathsf{T}}}^{[\mathsf{ci}]} = p_1 \wedge p_2 \wedge \Box((p_1 \wedge p_2 \wedge p \to \bot \vee \bot) \wedge p) \to \bot \vee \bot\ .$$

3. For the proto rule $\widehat{R_{45}}$ for the rule R_{45} given by the context $\Box p_1 \Rightarrow \Box q_1$ we get the substitution $\theta_{\widehat{R_{45}}}^{\mathsf{c}}$ with $\theta_{\widehat{R_{45}}}^{\mathsf{c}}(p) = (\Box p_1 \wedge \top \to \Box q_1 \vee p) \to p$ and $\theta_{\widehat{R_{45}}}^{\mathsf{c}}(p_1) = p_1$ and

$\theta^{\mathsf{c}}_{\widehat{R_{45}}}(q_1) = q_1$. Thus the proto rule $\widehat{R_{45}}$ is equivalent over Mon to the axiom

$$\mathsf{Concl}_{\widehat{R_{45}}}\theta^{\mathsf{c}}_{\widehat{R_{45}}} = \Box p_1 \wedge \top \to \Box q_1 \vee \Box((\Box p_1 \wedge \top \to \Box q_1 \vee p) \to p)\ .$$

Theorem 3.3.14 shows that for monotone modalities translatable clauses correspond very closely to proto rules for rules in normal form: every translatable clause is equivalent to a proto rule for a rule in normal form and every such proto rule is equivalent to a translatable clause. Thus from a point of view leaning more towards Hilbert systems it would be more natural to consider sequent calculi given by sets of proto rules and to consider rules with restrictions as particular sets of proto rules. Nevertheless, we are mainly interested in rules with restrictions and their corresponding axioms. By considering the equivalent sets of proto rules we obtain a correspondence here as well.

Theorem 3.3.16 (ci)**.** *Every rule R with finite restrictions in normal form is equivalent over* Mon *to an ω-set for a translatable clause A_R. If for two sets C_ℓ and C_r of formulae with $C_\ell \cap C_r = \emptyset$ all context restrictions of R have the form $\langle F_1, F_2\rangle$ with $F_1 \subseteq C_\ell$ and $F_2 \subseteq C_r$, then A_R has context formulae in (C_ℓ, C_r). In case all context restrictions of R are normal for a rule set $\mathcal{R}$ the rule R is equivalent over* Mon *to a single translatable clause A_R. If R is a shallow (resp. one-step) rule, then it is equivalent over* Mon *to a non-iterative (resp. rank-1) translatable clause.*

Proof. Let $R = \mathcal{P}/\Sigma \Rightarrow \Pi$ be a rule with finite context restrictions. It is clear that the rule R is equivalent to the set of proto rules for it. Furthermore by Theorem 3.3.14 every such proto rule is equivalent over Mon to a translatable clause. We would like to show that the set of all these clauses is axiomatically equivalent to an ω-set for a translatable clause. To construct this clause, note that since the restrictions of R are finite, we may simply take all formulae occurring in the restrictions as the context of a proto rule. More precisely, let $\widehat{R}$ be the canonical proto rule for R, i.e. the proto rule for R given by the context $\Gamma \Rightarrow \Delta$, where

$$\Gamma = \{C \mid C \in F_1 \text{ for some } \langle F_1, F_2\rangle \text{ occurring in } \mathcal{P}\}$$
$$\Delta = \{D \mid D \in F_2 \text{ for some } \langle F_1, F_2\rangle \text{ occurring in } \mathcal{P}\}\ .$$

Then by propositional reasoning the set of translations of proto rules for $\widehat{R}$ is axiomatically equivalent to the ω-set for the translation $A_{\widehat{R}}$ of the canonical proto rule $\widehat{R}$. Since $A_{\widehat{R}}$ has context formulae in $(\mathsf{Supp}(\Gamma), \mathsf{Supp}(\Delta))$ the claim about the context formulae follows. If all restrictions occurring in R are normal for a rule set $\mathcal{R}$, then by Lemma 3.2.22 the rule R is equivalent over Mon to the proto rule $\widehat{R}$ and thus also to the translatable clause $A_{\widehat{R}}$. This also holds for shallow or one-step rules, since the restrictions for such rules are always normal. Moreover, in these cases the resulting axioms are easily seen to be equivalent to non-iterative resp. rank-1 axioms. $\Box$

Example 3.3.17. 1. Using the method of the previous Theorem the set of axioms intuitionistically equivalent to the rule R_{4} is the ω-set for the translation of the canonical proto rule $\widehat{R_{\mathsf{4}}}$ for the rule R_{4} given by the context $\Box p_1 \Rightarrow$. This set has the form

$$\left\{ \bigwedge_{i=1}^{n} \Box p_i \wedge \top \to \bot \vee \Box \bigwedge_{i=1}^{n} \Box p_i \mid n \geq 0 \right\} .$$

By normality of $\Box p_1$ for $\mathcal{R}_{\mathsf{K}}$ this is equivalent over $\mathcal{R}_{\mathsf{K}}$ to the translation $\Box p_1 \wedge \top \to \bot \vee \Box\Box p_1$ of $\widehat{R_{\mathsf{4}}}$, which again is equivalent to the axiom $(\mathsf{4})$ $\Box p_1 \to \Box\Box p_1$.

2. Similarly, rule R_{T} is classically equivalent to the ω-set

$$\left\{ \bigwedge_{i=1}^{n} p_i \wedge \Box((\bigwedge_{i=1}^{n} p_i \wedge p \to \bigvee_{j=1}^{m} q_j \vee \bot) \wedge p) \to \bigvee_{j=1}^{m} q_j \vee \bot \mid n, m \geq 0 \right\}$$

for the translation of the canonical proto rule $\widehat{R_{\mathsf{T}}}$ for the rule R_{T} given by the context $p_1 \Rightarrow q_1$. In the intuitionistic case the index m ranges only over $0, 1$. By normality of propositional variables this is equivalent over $\mathcal{R}_{\mathsf{K}}$ to the translation $p_1 \wedge \Box((p_1 \wedge p \to q_1 \vee \bot) \wedge p) \to q_1 \vee \bot$ of $\widehat{R_{\mathsf{T}}}$, and by propositional reasoning and monotonicity of $\Box$ this is axiomatically equivalent to the axiom (T) $\Box s \to s$.

3. Finally, rule R_{45} is classically equivalent to the ω-set

$$\left\{ \bigwedge_{i=1}^{n} \Box p_i \wedge \top \to \bigvee_{j=1}^{m} \Box q_j \vee \Box((\bigwedge_{i=1}^{n} \Box p_i \wedge \top \to \bigvee_{j=1}^{m} \Box q_j \vee p) \to p) \mid n, m \geq 0 \right\}$$

for the translation of the canonical proto rule $\widehat{R_{\mathsf{45}}}$ for the rule R_{45} given by the context $\Box p_1 \Rightarrow \Box q_1$. Since the formula $\Box q_1$ is not right normal we cannot use Theorem 3.3.16 to turn this into a single axiom. But by propositional reasoning this set is axiomatically equivalent to the set consisting of axioms

$$\bigwedge_{i=1}^{n} \Box p_i \to \Box \left(\left(\bigwedge_{i=1}^{n} \Box p_i \wedge \neg \bigvee_{j=1}^{m} \Box q_j \right) \vee p \right) \vee \bigvee_{j=1}^{m} \Box q_j$$

for $n, m \geq 0$. By monotonicity of $\Box$ and propositional reasoning this set is moreover axiomatically equivalent to the set

$$\left\{ \bigwedge_{i=1}^{n} \Box p_i \to \Box \left(\bigwedge_{i=1}^{n} \Box p_i \wedge \neg \bigvee_{j=1}^{m} \Box q_j \right) \vee \bigvee_{j=1}^{m} \Box q_j \mid n, m \geq 0 \right\} ,$$

which by normality of $\Box$ and propositional reasoning is axiomatically equivalent to

$$\left\{\bigwedge_{i=1}^{n} \Box p_i \to \Box \bigwedge_{i=1}^{n} \Box p_i \mid n \geq 0\right\} \quad \cup \quad \left\{\neg \bigvee_{j=1}^{m} \Box q_j \to \Box\neg \bigvee_{j=1}^{m} \Box q_j \mid m \geq 0\right\}.$$

Here the first set is the ω-set for the axiom (4) $\Box p_1 \to \Box\Box p_1$ and by normality of $\Box$ axiomatically equivalent to it. The second set is the ω-set for the axiom (5) $\neg\Box q_1 \to \Box\neg\Box q_1$. In this particular case we are lucky and by normality of $\Box$ and propositional reasoning this set is equivalent to $\{\bigwedge_{j=1}^{m} \neg\Box q_j \to \bigwedge_{j=1}^{m} \Box\neg\Box q_j \mid m \geq 0\}$, which is axiomatically equivalent to the single axiom (5).

The results of this section together with those of the previous section allow us to characterise for monotone modalities the class of axioms which are equivalent to rules with restrictions purely in the Hilbert style setting.

Theorem 3.3.18 (ci)**.** *For monotone modalities an axiom for a Hilbert style system is*

1. *equivalent to a set of rules with restrictions iff axiomatically equivalent to the union of a set of ω-sets for translatable clauses*
2. *equivalent to a finite set of rules with finite restrictions iff axiomatically equivalent to the union of a finite set of ω-sets for translatable clauses*
3. *equivalent to a finite set of rules with finite normal restrictions iff axiomatically equivalent to a finite set of translatable clauses with normal context formulae*
4. *equivalent to a finite set of shallow rules iff equivalent to a finite set of non-iterative translatable clauses (in the classical case: iff equivalent to a non-iterative axiom)*
5. *equivalent to a finite set of one-step rules iff equivalent to a finite set of rank-1 translatable clauses (in the classical case: iff equivalent to a rank-1 axiom).*

Proof. Using the translations in Section 3.2 and Theorem 3.3.16 and the facts that by Lemma 3.3.4 every non-trivial rule with finite restrictions is equivalent to a finite set of rules in normal form and that by Lemma 3.3.1 every rule with restrictions is equivalent to a set of rules with finite restrictions. $\Box$

The correspondences are shown diagrammatically in Table 3.2.

3.4 Applications: Limitative Results

The correspondence results of the previous two sections can be applied in at least two ways. One possibility is to use the translation from axioms into rules to construct new sequent

set of ω-sets for translatable clauses	$\leftrightsquigarrow$	set of rules with restrictions
ω-set for a translatable clause	$\leftrightsquigarrow$	rule with finite restrictions in normal form
translatable clause with normal context formulae	$\leftrightsquigarrow$	rule with finite normal restrictions in normal form
translatable clause	$\leftrightsquigarrow$	proto rule for a rule with restrictions in normal form
non-iterative translatable clause	$\leftrightsquigarrow$	shallow rule
rank-1 translatable clause	$\leftrightsquigarrow$	one-step rule

Table 3.2: The corresponding classes of axioms and rules

calculi for modal logics given axiomatically in the spirit of Question 1.1.3. Of course this possibility hinges on methods to make the resulting sequent calculus cut-free. We will explore this issue further in the following chapters, where we will also see the translation in action for other than the standard examples. The second possibility is to follow Question 1.1.4 and show limitative results about which kinds of modal logics can be given a sound and complete sequent system with rules of a specific format. Such limitative results show on the one hand how much additional machinery beyond Gentzen's original rule format is necessary to capture a specific logic. On the other hand these results might also help in constructing sequent calculi for new modal logics, since they can be used to decide which kinds of calculi to search for. We will now see some results in this spirit, beginning with limitative results about the expressive strength of the different rule formats. It should be noted that since the translations make heavy use of the cut rule, the results limit the expressive strength of systems *with* the cut rule. Of course we may always add the cut rule to a cut-free system, and then apply the results to the new system. Here we only consider the classical setting. We will make use of the following standard notions for normal modal logics based on classical propositional logic, see e.g. [BdRV01].

Definition 3.4.1 (c)**.** A *Kripke frame* is a tuple $\mathfrak{F} = (W, R)$ consisting of a non-empty set W of *worlds* and binary *accessibility relation* $R \subseteq W \times W$ on the set of worlds. A *Kripke model* $(\mathfrak{F}, \sigma)$ is a Kripke frame $\mathfrak{F}$ together with a *valuation* $\sigma : W \to \mathfrak{P}(\mathsf{Var})$. The model $(\mathfrak{F}, \sigma)$ then is *based on* the frame $\mathfrak{F}$. We also write $\mathcal{F}_\Box$ for $\mathcal{F}(\{\wedge, \vee, \to, \bot, \Box\})$. If $\mathfrak{F} = (W, R)$ and $(\mathfrak{F}, \sigma)$ is a Kripke model, then *satisfaction* for a formula $A \in \mathcal{F}_\Box$ at a world w in this model is denoted by $\mathfrak{F}, w, \sigma \Vdash A$ and is recursively defined by

$$\mathfrak{F}, w, \sigma \Vdash p \text{ iff } p \in \sigma(w) \text{ for } p \in \mathsf{Var}$$
$$\mathfrak{F}, w, \sigma \Vdash \Box A \text{ iff for all } v \in W \text{ with } wRv \text{ we have } \mathfrak{F}, v, \sigma \Vdash A$$

and the standard clauses for the boolean connectives. A formula $A \in \mathcal{F}_\Box$ is *valid* in a model $(\mathfrak{F}, \sigma)$ if $\mathfrak{F}, w, \sigma \Vdash A$ for every world w of $\mathfrak{F}$. It is moreover *valid* in the frame $\mathfrak{F}$ if it is valid in every model based on the frame $\mathfrak{F}$. If a formula A is valid in a frame $\mathfrak{F}$ we also write $\mathfrak{F} \Vdash A$. For a set $\mathcal{A}$ of formulae we write $\mathfrak{F} \Vdash \mathcal{A}$ if $\mathfrak{F} \Vdash A$ for every $A \in \mathcal{A}$. Finally, a formula A is *satisfiable* in a frame $\mathfrak{F} = (W, R)$ if there are a world $w \in W$ and a valuation σ such that $\mathfrak{F}, w, \sigma \Vdash A$.

We graphically represent Kripke frames (W, R) in the standard way by drawing an arrow from world w to world v if wRv holds. A little thought shows that a formula A is valid in a frame $\mathfrak{F}$ if and only if its negation $\neg A$ is not satisfiable in $\mathfrak{F}$. In the following it will also be convenient to use $\Diamond A$ as an abbreviation for $\neg\Box\neg A$. We will be interested in classes of frames which are *defined* by a set of modal formulae.

Definition 3.4.2 (c)**.** Let F be a class of Kripke frames and let $\mathcal{A} \subseteq \mathcal{F}_\Box$ be a set of modal formulae. We say that $\mathcal{A}$ *modally defines* the class F if for every Kripke frame $\mathfrak{F}$ we have

$$\mathfrak{F} \in \mathsf{F} \Longleftrightarrow \mathfrak{F} \Vdash \mathcal{A} .$$

The class F is *modally definable* if it is modally defined by some set $\mathcal{A}$ of modal formulae. Furthermore, we write $\mathcal{L}_\mathsf{F}$ for the set $\{A \in \mathcal{F}_\Box \mid \mathfrak{F} \Vdash A \text{ for every } \mathfrak{F} \in \mathsf{F}\}$ of modal formulae valid in every frame of the class F. Given a set $\mathcal{A}$ of axioms and a class F of frames the Hilbert system $\mathcal{H}\mathsf{c}\mathcal{A}$ is *sound* for the class F if $\mathcal{H}\mathsf{c}\mathcal{A} \subseteq \mathcal{L}_\mathsf{F}$ and *complete* if $\mathcal{L}_\mathsf{F} \subseteq \mathcal{H}\mathsf{c}\mathcal{A}$.

We will make use of the following small lemma about alternative characterisations of modally definable frame classes from Hilbert-axiomatisations, where we write $\mathcal{H}\mathsf{cK}$ for the Hilbert system given by the set $\{\Box p \land \Box q \leftrightarrow \Box(p \land q),\ \Box\top\}$ of axioms and $\mathcal{H}\mathsf{cK}\mathcal{A}$ for the system given by these axioms together with the axioms in $\mathcal{A}$. It is well-known that $\mathcal{H}\mathsf{cK}$ is sound and complete for the class of all frames [BdRV01].

Lemma 3.4.3 (c)**.** *Let F be a modally definable class of Kripke frames and $\mathcal{A} \subseteq \mathcal{F}_\Box$ a set of modal formulae. If $\mathcal{H}\mathsf{cK}\mathcal{A}$ is sound and complete for F, then $\mathcal{A}$ modally defines the class F.*

Proof. Let F be modally defined by a set $\mathcal{B}$ of axioms. We need to show that for every frame $\mathfrak{F}$ we have $\mathfrak{F} \in \mathsf{F} \Leftrightarrow \mathfrak{F} \Vdash \mathcal{A}$. So take an arbitrary frame $\mathfrak{F}$.

Suppose that $\mathfrak{F} \in \mathsf{F}$. Since $\mathcal{H}\mathsf{cK}\mathcal{A}$ is sound for F we know that every formula derivable in $\mathcal{H}\mathsf{cK}\mathcal{A}$ is valid in every frame in F, and thus also in $\mathfrak{F}$. But since the axioms $\mathcal{A}$ obviously are derivable in $\mathcal{H}\mathsf{cK}\mathcal{A}$ we thus have $\mathfrak{F} \Vdash \mathcal{A}$.

Now suppose that $\mathfrak{F} \Vdash \mathcal{A}$. Since $\mathcal{H}\mathsf{cK}\mathcal{A}$ is complete for F we know that every formula valid in every frame in F is derivable in $\mathcal{H}\mathsf{cK}\mathcal{A}$. Thus, since $\mathcal{B}$ modally defines F in particular every formula in $\mathcal{B}$ is derivable in $\mathcal{H}\mathsf{cK}\mathcal{A}$. But since $\mathfrak{F} \Vdash \mathcal{A}$ we also know that every formula derivable in $\mathcal{H}\mathsf{cK}\mathcal{A}$ must be valid in $\mathfrak{F}$, and thus we get $\mathfrak{F} \Vdash \mathcal{B}$. But $\mathcal{B}$ modally defines F, and thus we have $\mathfrak{F} \in \mathsf{F}$. $\square$

Remark 3.4.4. Note that the condition of modal definability of the class F in Lemma 3.4.3 is crucial: it is well-known that the Hilbert-system $\mathcal{H}\mathsf{cK}$ is sound and complete for the class of *irreflexive* frames, a frame class which is not modally definable at all, see e.g. [HC96, p.176].

Using the previous Lemma it is possible to show that the Hilbert system generated by a set $\mathcal{A}$ of axioms is *not* sound and complete for a modally definable class of Kripke frames - we simply show that the set of axioms cannot possibly modally define the class of frames. For convenience we capture the general idea in the following Lemma.

Lemma 3.4.5 (c)**.** *Let* F *be a modally definable class of Kripke frames and let* $\mathcal{A} \subseteq \mathcal{F}_\Box$ *be a set of formulae. If there are two frames* $\mathfrak{F}_1, \mathfrak{F}_2$ *such that*

1. $\mathfrak{F}_1 \in \mathsf{F}$ *and* $\mathfrak{F}_2 \notin \mathsf{F}$
2. *for every formula* $A \in \mathcal{A}$ *we have* $\mathfrak{F}_1 \Vdash A \Leftrightarrow \mathfrak{F}_2 \Vdash A$,

then the logic $\mathcal{H}\mathsf{cK}\mathcal{A}$ *is not sound and complete for* F.

Proof. For the sake of contradiction, suppose that $\mathcal{H}\mathsf{cK}\mathcal{A}$ is sound and complete for F. Then by Lemma 3.4.3 we know that $\mathcal{A}$ modally defines F. But then by assumption 1 we have $\mathfrak{F}_1 \Vdash \mathcal{A}$, and thus by assumption 2 we also have $\mathfrak{F}_2 \Vdash \mathcal{A}$ in contradiction to the fact that $\mathfrak{F}_2 \notin \mathsf{F}$. $\square$

Thus in order to show that sequent systems including only rules of a certain format are not expressive enough to capture a particular axiom or modally definable logic it is enough to show that translations of such rules are not strong enough to define the class of frames for this logic. In the following we will make extensive use of this technique. The first goal is to show that the containments for the classes of logics which can be captured by one-step rules, shallow rules and rules with restrictions are proper. For this we will show the intuitively obvious results that one-step rules are not strong enough to capture reflexivity of the accessibility relation, and that shallow rules are not capable of capturing transitivity or symmetry.

Theorem 3.4.6 (c)**.** *There is no set of one-step rules equivalent over* $\mathcal{R}_\mathsf{K}$ *to the axiom* (T) $\Box p \to p$.

Proof. We know that the axiom (T) $\Box p \to p$ modally defines the class of Kripke frames with a reflexive accessibility relation [BdRV01]. On the other hand by Theorem 3.3.16 we know that the translations of one-step rules are rank-1 axioms. To see that no set of rank-1 axioms can be sound and complete for the class of transitive frames consider the two frames

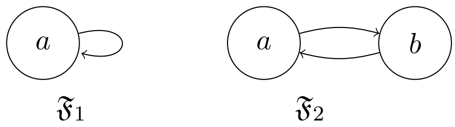


Figure 3.2: The two frames used in the proof of Theorem 3.4.6 (reflexivity).

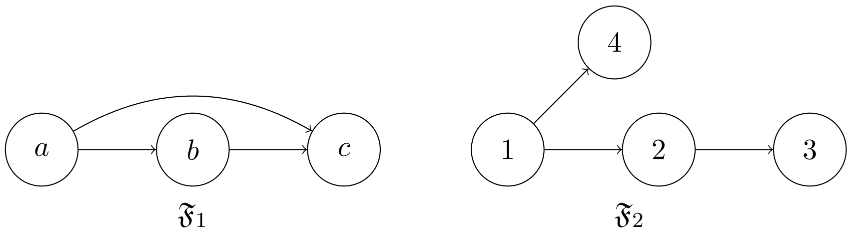


Figure 3.3: The two frames used in the proof of Theorem 3.4.7 (transitivity)

$\mathfrak{F}_1 := (\{a\}, \{(a,a)\})$ and $\mathfrak{F}_2 := (\{a,b\}, \{(a,b),(b,a)\})$ as shown in Figure 3.2. We will show that for every rank-1 axiom A we have $\mathfrak{F}_1 \Vdash A \Leftrightarrow \mathfrak{F}_2 \Vdash A$. The claim then follows immediately from Lemma 3.4.5. So let A be a rank-1 formula. We show the contrapositions of the above statement, that is we show that $\neg A$ is satisfiable in $\mathfrak{F}_1$ if and only if it is satisfiable in $\mathfrak{F}_2$. Obviously the negation of a rank-1 formula is rank-1 as well. Let σ_1 be a valuation such that $(\mathfrak{F}_1, a, \sigma_1) \Vdash \neg A$. Then for the valuation σ_2 on $\mathfrak{F}_2$ defined by $\sigma_2(a) = \sigma_2(b) = \sigma_1(a)$ we obviously have $(\mathfrak{F}_2, a, \sigma_2) \Vdash \neg A$ as well. For the other direction suppose that w.l.o.g. for a world a and a valuation σ_2 on $\mathfrak{F}_2$ we have $(\mathfrak{F}_2, a, \sigma_2) \Vdash \neg A$. Then setting $\sigma_1(a) := \sigma_2(b)$ and using the fact that since $\neg A$ is a rank-1 formula every propositional variable in $\neg A$ occurs under exactly one modality we obtain $(\mathfrak{F}_1, a, \sigma_1) \Vdash \neg A$ as well. $\square$

Thus, since the axiom (T) is equivalent to a shallow rule (see Example 3.2.25), the class of modal logics that can be captured by one-step rules is properly contained in the class of modal logics that can be captured by shallow rules.

Theorem 3.4.7 (c)**.** *There is no set of shallow rules equivalent over $\mathcal{R}_{\mathsf{K}}$ to either of the axioms* (4) $\Box\Box p \to \Box p$ *or* (B) $p \to \Box\Diamond p$.

Proof. We use the same technique as above, making use of the fact that by Theorem 3.3.16 translations of shallow rules are non-iterative formulae.

The axiom (4) $\Box\Box p \to \Box p$ modally defines the class of Kripke frames where the accessibility relation is transitive [BdRV01]. Consider the two frames $\mathfrak{F}_1 := (\{a,b,c\}, \{(a,b),(b,c),(a,c)\})$ and $\mathfrak{F}_2 := (\{1,2,3,4\}, \{(1,2),(2,3),(1,4)\})$ as shown in Figure 3.3. Now suppose that A is a non-iterative formula. Then $\neg A$ is non-iterative as well. Again we show that $\neg A$ is satisfiable in $\mathfrak{F}_1$ if and only if it is satisfiable in $\mathfrak{F}_2$. Suppose we have a valuation σ_2 on $\mathfrak{F}_2$ and a world w_2 with $(\mathfrak{F}_1, w_2, \sigma_2) \Vdash \neg A$. Define a valuation σ_1 on $\mathfrak{F}_1$ and a world w_1 as follows:

- if $w_2 = 1$, then set $\sigma_1(a) := \sigma_2(1), \sigma_1(b) := \sigma_2(2), \sigma_1(c) := \sigma_2(4)$ and set $w_1 := a$
- if $w_2 = 2$, then set $\sigma_1(b) := \sigma_2(2), \sigma_1(c) := \sigma_2(3)$ and $\sigma_1(a)$ arbitrary, and set $w_1 := b$
- if $w_2 \in \{3, 4\}$, then set $\sigma_1(c) := \sigma_2(w_2)$ and $\sigma_1(a)$ and $\sigma_1(b)$ arbitrary, and set $w_1 := c$.

Then it is not hard to see that $(\mathfrak{F}_1, \sigma_1, w_1) \Vdash \neg A$. The other direction is analogous but easier. Thus non-iterative axioms cannot define the class of transitive frames and we are done using Lemma 3.4.5.

The proof for the axiom (B) $p \to \Box\Diamond p$ is analogous using the two frames $\mathfrak{F}_1 := (\mathbb{N}, \mathsf{succ})$ and $\mathfrak{F}_2 := (\{a, b\}, \{(a, b), (b, a)\})$ and the fact that (B) modally defines the class of Kripke frames where the accessibility relation is symmetric [BdRV01]. $\Box$

Since the axiom $(\mathsf{4})$ is equivalent to a rule with restrictions (again see Example 3.2.25), this establishes that the class of logics which can be captured by shallow rules is properly contained in the class of logics which can be captured by rules with restrictions.

Corollary 3.4.8 (c)**.** *Let L_1 (resp. L_{0-1} resp. L) be the class of logics for which there is a sound and complete (with GcCut) sequent system given by a set of modal one-step rules (resp. shallow rules resp. rules with restrictions). Then $\mathsf{L}_1 \subsetneqq \mathsf{L}_{0-1} \subsetneqq \mathsf{L}$.* $\Box$

We can also use Lemma 3.4.3 in a slightly different way to establish impossibility results by making use of the following result.

Definition 3.4.9. The *first-order correspondence language* is the first-oder language containing equality and a binary relation symbol. A class F of frames is *first-order definable* if there is a formula φ in the first-order correspondence language such that for every frame $\mathfrak{F}$ we have $\mathfrak{F} \in \mathsf{F}$ iff $\mathfrak{F} \models \varphi$ when interpreted as a first-order structure. A modal formula $A \in \mathcal{F}_\Box$ is *elementary* if the class F of frames modally defined by A is first-order definable.

Theorem 3.4.10 ([vB83, tC05])(c)**.** *Every non-iterative modal formula is elementary.*

Corollary 3.4.11 (c)**.** *There is no set of shallow rules equivalent over $\mathcal{R}_\mathsf{K}$ to either of the axioms* (L) $\Box(\Box p \to p) \to \Box p$ *or* (Grz) $\Box(\Box(p \to \Box p) \to p) \to p$.

Proof. The axiom (L) $\Box(\Box p \to p) \to \Box p$ modally defines the class of Kripke frames where the accessibility relation is transitive and conversely well founded, and the axiom (Grz) $\Box(\Box(p \to \Box p) \to p) \to p$ modally defines the class of frames where the accessibility relation R is reflexive and transitive and where there are no infinite paths $x_0 R x_1 R x_2 R \dots$ with $x_i \neq x_{i+1}$ for every $i \geq 0$. Moreover, it is known that neither of these classes of frames is first-order definable [BdRV01]. But if either of these classes were captured by a set of shallow rules, then by the translation of Theorem 3.3.16 and by Lemma 3.4.3 it would be modally defined by a set of non-iterative axioms, and therefore by Theorem 3.4.10 and the results in [vB76] it would be first-order definable. Thus there are no such rule sets. $\Box$

So far the results have only concerned the rather restrictive rule formats of one-step rules and shallow rules. Unfortunately it is not clear whether the technique of Lemma 3.4.5 can be applied to rules with restrictions in general - the format of a translatable axiom might be to general for this. But we might try to impose some extra conditions on the rule sets in order to suitably restrict the format of the resulting axioms. Since we are in general mainly interested in finding cut-free sequent systems, it might be tempting to simply stipulate that the sequent system generated by our rules be cut-free. But unfortunately this is slightly problematic, since our translation makes essential use of the cut rule. So we need to impose some further restrictions. Here we are going to consider the restriction that the rule set is *mixed-cut closed* in the sense of Definition 2.4.11. As we have seen in Corollary 2.5.7 this condition entails a more restricted format of the rules. We will see that this restricted format enables us to show a number of limitative results.

The choice of the restriction to mixed-cut closed rule sets warrants some discussion. While mixed-cut closure is a property of *rule sets*, this property is suggested by considering *proofs* of syntactic cut elimination, and in particular proofs which are 'reasonably standard'. For the latter it seems fair to assume that a 'reasonably standard' proof proceeds in the spirit of Gentzen's original proof and makes essential use of a permutability-of-rules argument. In particular, one of the main ingredients of such a proof seems to be the technique of permuting a cut on a modal formula, which is principal in the last applied rule on one side and contextual in the last applied rule on the other side, into the premisses of the latter rule. But stipulating that it is always possible to do so amounts to stipulating that the set of modal rules is mixed-cut closed. It seems that as far as only modal rules are concerned this permutation argument is essential in virtually all proofs of cut elimination which rely on a modification of Gentzen's original proof. Moreover, in the context of *standard calculi* the closely related condition of *weak substitutivity* has been shown to be necessary for a strong form of cut elimination [CT06a, CT06b]. In case the cut formula is a propositional formula, the other main technique is to use the Inversion Lemma for the propositional connectives. But this lemma also usually relies on a permutation-of-rules argument, which amounts to stipulating Gc-*invertibility* in the sense of Definition 2.5.3. Thus it seems natural to assume that a reasonably standard proof for cut-elimination for a sequent system given by $\mathsf{Gc}\mathcal{R}$ makes use of mixed-cut closure of $\mathcal{R}$ and either mixed-cut closure of $\mathsf{Gc}\mathcal{R}$ or Gc-invertibility for $\mathcal{R}$. Now in view of Theorem 2.5.6 whenever we have a mixed-cut closed rule set $\mathcal{R}$ which is Gc-inverting we can find an equivalent rule set $\mathcal{R}'$ such that $\mathsf{Gc}\mathcal{R}'$ is mixed-cut closed. Thus we might equivalently stipulate that our rules including the propositional rules are mixed-cut closed.

It is important to note that even though the restriction to mixed-cut closed rule sets is suggested by the standard (syntactical) technique for proving cut elimination, the results themselves only rely on a property of the rule set. Thus they are independent of any particular proof technique and preclude also e.g. semantical proofs of cut admissibility. In view of

Theorem 2.5.6 we only state these results with respect to rule sets $\mathcal{R}$ where $\mathsf{Gc}\mathcal{R}$ is mixed-cut closed, but the reader should bear in mind that they hold for rule sets where $\mathcal{R}$ is mixed-cut closed and Gc-permuting as well. The first step is to translate the restricted rule format from Corollary 2.5.7 to the Hilbert-setting. Recall that a logic $\mathcal{L}$ has *non-trivial modalities* if for every modality neither $\models_{\mathcal{L}} \heartsuit\vec{p}$ nor $\models_{\mathcal{L}} \neg\heartsuit\vec{p}$ (see Theorem 2.5.6).

Lemma 3.4.12 (c)**.** *Let $\mathcal{L}$ be logic with non-trivial modalities and* monotonicity*, i.e. whenever $\models_{\mathcal{L}} p_i \to q_i$ for $i \leq n$, then $\models_{\mathcal{L}} \heartsuit(p_1, \dots, p_n) \to \heartsuit(q_1, \dots, q_n)$. If there is a set $\mathcal{R}$ of modal rules (not necessarily including* Cong *or* Mon*) such that* $\mathsf{Gc}\mathcal{R}$ *is mixed-cut closed and* $\mathsf{Gc}\mathcal{R}\mathsf{ConW}$ *is sound and complete for $\mathcal{L}$, then $\mathcal{L}$ can be axiomatised in a Hilbert system by translatable clauses with context formulae in $(\Lambda(V_1) \cup V_2, \Lambda(V_3) \cup V_4)$ for $V_1, \dots, V_4 \subseteq \mathsf{Var}$ pairwise disjoint. If $\mathcal{L}$ is a normal $\Lambda_\Box$-logic, then the context formulae can be taken to be in $(\{p_1, \Box p_2\}, \{q_1, \Box q_2\})$.*

Proof. It is clear that if a logic satisfies monotonicity, then it also satisfies congruence. Now suppose we have such a rule set $\mathcal{R}$. Then by Corollary 2.5.7 w.l.o.g. every restriction occurring in a rule in $\mathcal{R}$ contains only variables or modalised variables. Furthermore, by renaming the variables we may assume w.l.o.g. that there are sets C_ℓ, C_r of formulae with $C_\ell \cap C_r$ such that for every restriction $\langle F_1; F_2\rangle$ of R we have $F_1 \subseteq C_\ell$ and $F_2 \subseteq C_r$. Since $\mathcal{L}$ has monotonicity the rules in Mon are sound and we may simply add them to the rule set. Doing this and then bringing the rules into standard form as in Lemma 3.3.4 does not change the form of the context formulae. Now translating the rules into equivalent axioms using Theorem 3.3.16 yields the result. The statement for normal modal logics follows immediately. ◻

This again gives us a restricted format for the axioms which we can use in combination with Lemma 3.4.5 to show impossibility results. In particular axioms of the form specified in the Lemma have modal nesting depth at most two, and the modalised context formulae use different variables from those appearing on the top level of the formula. The first property is already enough to give us an impossibility result for the logic of *2-transitive* frames.

Definition 3.4.13. A Kripke frame $\mathfrak{F} = (F, R)$ is *2-transitive* if $\mathfrak{F} \Vdash \Box\Box p \to \Box\Box\Box p$.

It is not too hard to see that the class of 2-transitive frames is first order defined by the formula

$$\forall w \forall x \forall y \forall z (wRx \wedge xRy \wedge yRz \to \exists v (wRv \wedge vRz)) \; .$$

While the axiom $\Box\Box p \to \Box\Box\Box p$ can easily be translated into the equivalent rule with restrictions $\{(\Rightarrow q; \langle\{\Box\Box p\}, \emptyset\rangle)\}/ \Rightarrow \Box q$ and thus is captured (with Cut) by a sequent system of rules with restrictions, it cannot be characterised by modal axioms of rank 2, and thus it cannot be captured by a mixed-cut closed set of rules with restrictions.

Theorem 3.4.14 (c)**.** *There is no rule set $\mathcal{R}$ (not necessarily containing* Cong *or* Mon*) such that* $\mathsf{Gc}\mathcal{R}$ *is mixed-cut closed and such that* $\mathsf{Gc}\mathcal{R}\mathsf{ConW}$ *is sound and (cut-free) complete for the logic of 2-transitive frames.*

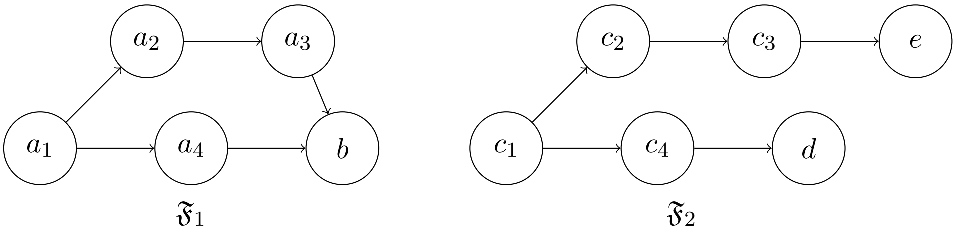


Figure 3.4: The two frames used in the proof of Theorem 3.4.14 (2-transitivity).

Proof. We use Lemma 3.4.5 to show that modal axioms of rank 2 cannot axiomatise this logic. The result then follows from Lemma 3.4.12.

So consider the two frames $\mathfrak{F}_1 = (\{a_1, a_2, a_3, a_4, b\}, \{(a_1, a_2), (a_2, a_3), (a_1, a_4), (a_3, b), (a_4, b)\})$ and $\mathfrak{F}_2 = (\{c_1, c_2, c_3, c_4, d, e\}, \{(c_1, c_2), (c_2, c_3), (c_1, c_4), (c_3, d), (c_4, e)\})$ as shown in Figure 3.4. The frame $\mathfrak{F}_1$ is 2-transitive, while the frame $\mathfrak{F}_2$ is not. Now let A be a modal formula with modal rank at most 2. Then clearly the formula $\neg A$ has modal rank at most 2 as well. But if there are a valuation σ_2 on $\mathfrak{F}_2$ and a world w_2 such that $\mathfrak{F}_2, w_2, \sigma_2 \Vdash \neg A$, then we construct a valuation σ_1 on $\mathfrak{F}_1$ and a world w_1 as follows: if for $i \in \{1, \dots, 4\}$ we have $w_2 = c_i$, then set $w_1 := a_i$; if $w_2 \in \{d, e\}$ set $w_1 := b$. For $i \in \{1, \dots, 4\}$ set $\sigma_1(a_i) := \sigma_2(c_i)$. If $w_2 \in \{c_1, c_4, d\}$ set $\sigma_1(b) := \sigma_2(d)$, otherwise set $\sigma_1(b) := \sigma_2(e)$. Then since $\neg A$ has modal rank at most 2 it is not hard to see that $\mathfrak{F}_1, w_1, \sigma_1 \Vdash \neg A$ as well. Similarly if for a valuation σ_1 and world w_1 we have $\mathfrak{F}_1, w_1, \sigma_1 \Vdash \neg A$ then we can construct a valuation σ_2 and a world w_2 with $\mathfrak{F}_2, w_2, \sigma_2 \Vdash \neg A$. Thus we have $\mathfrak{F}_1 \Vdash A$ iff $\mathfrak{F}_1 \Vdash A$ and Lemma 3.4.5 yields the result. □

The fact that translations of rules of this restricted form have modal rank 2 may already suffice to show some limitative results, but the second property of such translations is quite powerful as well: since in a rule with restrictions the variables occurring in a context formulae *only* occur in this formula, intuitively the context formulae cannot interact with other context formulae or the principal formulae. We use this to show that mixed-cut closed sets of rules with context restrictions cannot capture symmetry of the accessibility relation.

Theorem 3.4.15 (c)**.** *There is no rule set $\mathcal{R}$ (not necessarily containing* Cong *or* Mon*) such that* $\mathsf{Gc}\mathcal{R}$ *is mixed-cut closed and such that* $\mathsf{Gc}\mathcal{R}\mathsf{ConW}$ *is sound and (cut-free) complete for the logic of symmetric frames.*

Proof. Suppose that there is such a rule set. Then by Lemma 3.4.12 the logic of symmetric frames can be axiomatised by a set $\mathcal{A}$ of translatable clauses with context formulae in $(\{p_1, \Box p_2\}, \{q_1, \Box q_2\})$. Thus in formulae from $\mathcal{A}$ the variables p_2 and q_2 occur only boxed and every variable occurring under two layers of modalities is one of these two variables. Now consider the two frames $\mathfrak{F}_1 := (\{v, w\}, \{(v, w), (w, v)\})$ and $\mathfrak{F}_2 := (\mathbb{N}, \{(n, n + 1) \mid n \in \mathbb{N}\})$ as shown in Figure 3.5. Clearly $\mathfrak{F}_1$ is symmetric, while $\mathfrak{F}_2$ is not. We show that axioms in $\mathcal{A}$ are valid in $\mathfrak{F}_1$ iff they are valid in $\mathfrak{F}_2$. Suppose we have a formula $A \in \mathcal{A}$. Then clearly for $\neg A$ as

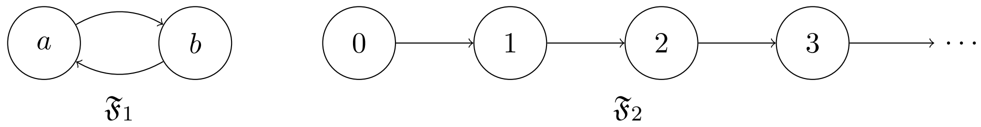


Figure 3.5: The two frames used in the proof of Theorem 3.4.15 (symmetry).

well the variables in $\{p_2, q_2\}$ only occur boxed and this set contains every variable occurring under two layers of modalities. Now suppose that for a world $w_2 \in \mathbb{N}$ and a valuation σ_2 on $\mathfrak{F}_2$ we have $\mathfrak{F}_2, w_2, \sigma_2 \Vdash \neg A$. W.l.o.g. we have $w_2 = 0$ and define a valuation σ_1 on $\{a, b\}$ by

$$\begin{aligned} \sigma_1 \restriction_{\{p_2,q_2\}} (a) &:= \sigma_2 \restriction_{\{p_2,q_2\}} (2) \\ \sigma_1 \restriction_{\mathsf{var}(A) \smallsetminus \{p_2,q_2\}} (a) &:= \sigma_2 \restriction_{\mathsf{var}(A) \smallsetminus \{p_2,q_2\}} (0) \\ \sigma_1(b) &:= \sigma_2(1) \end{aligned}$$

Then for the variables $p \in \{p_2, q_2\}$ we have $(\mathfrak{F}_1, a, \sigma_1 \Vdash \Box p$ iff $\mathfrak{F}_2, 0, \sigma_2 \Vdash \Box p)$ and $(\mathfrak{F}_1, b, \sigma_1 \Vdash \Box p$ iff $\mathfrak{F}_2, 1, \sigma_2 \Vdash \Box p)$. Also for variables $q \in \mathsf{var}(A) \smallsetminus \{p_2, q_2\}$ we have $(\mathfrak{F}_1, a, \sigma_1 \Vdash q$ iff $\mathfrak{F}_2, 0, \sigma_2 \Vdash q)$ and $(\mathfrak{F}_1, b, \sigma_1 \Vdash q$ iff $\mathfrak{F}_2, 1, \sigma_2 \Vdash q)$. Thus we have $\mathfrak{F}_1, a, \sigma_1 \Vdash \neg A$. Analogously if for a valuation σ_1 we have $\mathfrak{F}_1, a, \sigma_1 \Vdash \neg A$, then setting $\sigma_2(0) := \sigma_2(2) := \sigma_1(a)$ and $\sigma_2(1) := \sigma_1(b)$ we get $\mathfrak{F}_2, 0, \sigma_2 \Vdash \neg A$. Thus axioms in $\mathcal{A}$ are valid in $\mathfrak{F}_1$ iff they are valid in $\mathfrak{F}_2$ and an application of Lemma 3.4.5 together with the fact that the axiom (B) $p \to \Box\Diamond p$ modally defines the class of symmetric Kripke frames [BdRV01] yields a contradiction. $\Box$

Unlike in the case of 2-transitivity there is no known sequent system consisting of rules with restrictions which is sound and complete (with Cut) for the logic of symmetric frames. It is not clear whether the technique of this section can be adapted to show that there cannot be such a system. We encounter a similar situation in the case of the logic GL of transitive Kripke frames with a conversely well-founded accessibility relation. It is well-known that this class of frames is modally defined by the axiom (L) $\Box(\Box p \to p) \to \Box p$ and that its logic is axiomatised by adding (L) to an axiomatisation for the standard modal logic K (see e.g. [CZ97, BdRV01]). Again it is not clear whether this logic can be captured (using Cut) by rules with restrictions, but we can show that it cannot be captured by a mixed-cut closed set of rules with restrictions.

Theorem 3.4.16 (c)**.** *There is no rule set $\mathcal{R}$ (not necessarily containing* Cong *or* Mon*) such that* $\mathsf{G}c\mathcal{R}$ *is mixed-cut closed and such that* $\mathsf{G}c\mathcal{R}\mathsf{ConW}$ *is sound and (cut-free) complete for the logic* GL*.*

Proof. Again, if there were such a rule set, then by Lemma 3.4.12 this logic could be axiomatised by a set $\mathcal{A}$ of translatable clauses with context formulae in $(\{p_1, \Box p_2\}, \{q_1, \Box q_2\})$.

Consider the two transitive frames given by the infinite rooted trees with roots v resp.

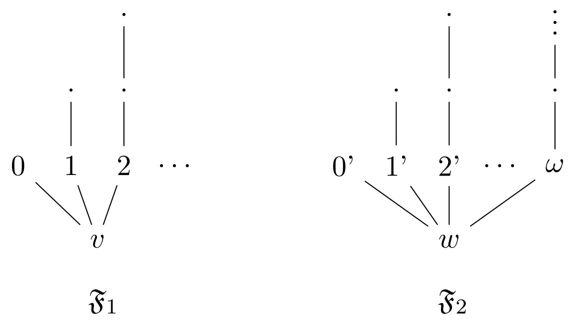


Figure 3.6: The two frames used in the proof of Theorem 3.4.16 (GL).

w in Figure 3.6 by setting xRy if x is on every path from the root to y and $x \neq y$. Since every branch in $\mathfrak{F}_1$ is finite the accessibility relation in $\mathfrak{F}_1$ is conversely well-founded and we have $\mathfrak{F}_1 \models \mathsf{GL}$. On the other hand, since $\mathfrak{F}_2$ contains an infinite branch we have $\mathfrak{F}_2 \not\models \mathsf{GL}$. For $i = 1, 2$ and $x \in F_i$ we write $\mathfrak{F}_i[x]$ for the subframe of $\mathfrak{F}_i$ generated by the set $\{y \mid x \text{ lies on every path from the root to } y\}$. Furthermore for $n \in \mathbb{N}$ we take f_n to be the canonical isomorphism $\mathfrak{F}_1[n] \to \mathfrak{F}_2[n']$. Now let A be an axiom of the form given above. Then

$$\neg A \;\equiv\; p_1 \wedge \Box p_2 \wedge P \wedge \bigwedge_{i \in I} \Box C_i \wedge \bigwedge_{j \in J} \Diamond \neg D_j \wedge \Diamond \neg q_2 \wedge \neg q_1$$

where for a set V of variables with $p_1, p_2, q_1, q_2 \notin V$ we have $P \in \mathsf{Prop}(V)$ and $C_i \in \mathcal{F}_r{}^{\mathsf{c}}(\{p_1, \Box p_2\}, V, \{q_1, \Box q_2\})$ and $D_j \in \mathcal{F}_\ell{}^{\mathsf{c}}(\{p_1, \Box p_2\}, V, \{q_1, \Box q_2\})$. Strictly speaking the context formulae $p_1, \Box p_2, \neg q_1, \Diamond \neg q_2$ need not appear in $\neg A$, but since this only makes the proof simpler we assume that they do appear. Thus w.l.o.g. the C_i have the form $p_1 \wedge \Box p_2 \wedge P' \to q_1 \vee \Box q_2$ and the $\neg D_j$ have the form $\neg(q_1 \wedge \Box q_2 \wedge P'' \to p_1 \vee \Box p_2) \equiv q_1 \wedge \Box q_2 \wedge P'' \wedge \neg p_1 \wedge \Diamond \neg p_2$, where $P', P'' \in \mathsf{Prop}(V)$ and where again the formulae $p_1, \Box p_2, q_1, \Box q_2$ need not occur, but if they occur, then they also occur on the top level of $\neg A$. Now suppose there are a valuation σ_1 and $y \in F_1$ such that $\mathfrak{F}_1, \sigma_1, y \Vdash \neg A$.

If $y \neq v$, then there is a $n \in \mathbb{N}$ with $y \in \mathfrak{F}_1[n]$ and for $z \in \mathfrak{F}_2[n']$ we set $\sigma_2(z) := \sigma_1(f_n^{-1}(z))$. Then $\mathfrak{F}_2, \sigma_2, f_n(y) \Vdash \neg A$.

If $y = v$, then for $z \in F_2$ we set

$$\sigma_2(z) \;:=\; \begin{cases} \sigma_1(f_n^{-1}(z)) & : \quad z \in \mathfrak{F}_2[n'] \\ \{p_2, q_1, q_2\} \cup (\sigma_1(0) \cap V) & : \quad z \in \mathfrak{F}_2[\omega] \\ \sigma_1(v) & : \quad z = w \end{cases}$$

Then for every $z \in \mathfrak{F}_2[\omega]$ and for every $i \in I$ we have $\mathfrak{F}_2, \sigma_2, z \Vdash C_i$ iff $\mathfrak{F}_1, \sigma_1, 0 \Vdash C_i$. Thus for every such z we have $\mathfrak{F}_2, \sigma_2, z \Vdash \bigwedge_{i \in I} C_i$ and thus since $\neg q_2$ and every $\neg D_j$ are already satisfied in a $\mathfrak{F}_2[n']$ we have $\mathfrak{F}_2, \sigma_2, w \Vdash \neg A$.

For the other direction suppose there are a valuation σ_2 and $z \in F_2$ such that $\mathfrak{F}_2, \sigma_2, z \Vdash \neg A$. If there is a $n \in \mathbb{N}$ with $z \in \mathfrak{F}_2[n']$, then similar to above for $y \in \mathfrak{F}_1[n]$ we set $\sigma_1(y) := \sigma_2(f_n(y))$ and obtain $\mathfrak{F}_1, \sigma_1, f_n^{-1}(z) \Vdash \neg A$.

If $z \in \mathfrak{F}_2[\omega]$, then for $j \in J$ let y_{D_j} and y_{q_2} be the first points each in $\mathfrak{F}[\omega]$ satisfying $\neg D_j$ respectively $\neg q_2$, and let m be the maximal distance of either of these from z. Note that in this case none of the $\neg D_j$ can contain the formula $\Diamond\neg p_2$, since then the formula $\Box p_2$ would need to appear on the top level of $\neg A$ as well and using transitivity we would have the contradiction $\mathfrak{F}_2, z, \sigma_2 \Vdash \Box p_2 \wedge \Diamond\neg p_2$. Now it is enough to copy the valuations of points in the initial part F of $\mathfrak{F}_2[z]$ of length m to points in $\mathfrak{F}_1[m+1]$ by using the canonical isomorphism $f : \mathfrak{F}_1[m+1] \to F$ and setting $\sigma_1(y) := \sigma_2(f(y))$ for $y \in \mathfrak{F}_1[m+1]$. This gives $\mathfrak{F}_1, \sigma_1, m+1 \Vdash \neg A$.

If $z = w$, then similarly to the last case we take y_{D_j} and y_{q_2} to be the first points each in $\mathfrak{F}_2[\omega]$ to satisfy a $\neg D_j$ resp. $\neg q_2$ which is not already satisfied in $\mathfrak{F}_2[n']$ for any $n \in \mathbb{N}$. Then we copy the valuations of points in the initial part of $\mathfrak{F}_2[\omega]$ of sufficient length to a branch $\mathfrak{F}_1[m]$ of $\mathfrak{F}_1$ where $\mathfrak{F}_2[m']$ is not the only branch witnessing any of the $\neg D_j$ or $\neg q_2$ and m is at least the maximal distance from ω to any of the y_{D_j}, y_{q_2}. For every $y \in \mathfrak{F}_1[n]$ with $n \neq m$ we set $\sigma_1(y) := \sigma_2(f_n(y))$, and $\sigma_1(v) := \sigma_2(w)$. Then every $\neg D_j$ is witnessed in a $\mathfrak{F}_1[n]$ and thus we have $\mathfrak{F}_1, \sigma_1, v \Vdash \neg A$.

Thus in total we get $\mathfrak{F}_1 \Vdash A$ iff $\mathfrak{F}_2 \Vdash A$. Now Lemma 3.4.5 gives the result. $\Box$

In some cases it is possible to use additional properties of the logics to further restrict the format of the sequent rules and with this the format of the axioms. As an example we consider the modal logic S5. This logic is particularly interesting, since even though it has a relatively simple semantic characterisation as the modal logic of Kripke frames where the accessibility relation is an equivalence relation [BdRV01] and even though as we have seen earlier it is captured by a sequent system with Cut and rules with restrictions (Example 2.3.7), so far it has eluded all efforts to construct a "standard" cut-free sequent system for it. Of course there are cut-free sequent systems for S5 in extended sequent frameworks such as hypersequents [Avr96, Pog08], nested sequents [Brü09] or labelled calculi [Neg05], but all of these require additional machinery beyond the standard framework of two-sided sequents. The following results might be seen as a step towards a formal explication why finding a "standard" cut-free system is very hard in the least. As a first step since the formula $\Box p \to p$ must be derivable in a cut-free sequent system for S5 we can strengthen the claim of Corollary 2.5.7.

Lemma 3.4.17 (c)**.** *Let $\mathcal{R}$ be a set of modal rules with restrictions (not necessarily including* Cong *or* Mon*) such that* Gc$\mathcal{R}$ConW *is sound and cut-free complete for* S5 *and such that* Gc$\mathcal{R}$ *is mixed-cut closed. Then w.l.o.g. for every restriction $\langle F_1, F_2\rangle$ of a rule in $\mathcal{R}$ we have $F_1 \subseteq \{\Box p, p\}$ and $F_2 \subseteq \{p\}$.*

Proof. Similar to the proof of Theorem 2.5.6 we first show that by mixed-cut closure of Gc$\mathcal{R}$ whenever a rule has a restriction $\langle F_1, F_2\rangle$ such that $p \in F_1$ (resp. $p \in F_2$) for a variable p, then also $p \in F_2$ (resp. $p \in F_1$). For this again we apply the condition of mixed-cut closure to the rule in question and one of the rules $\vee_{\mathsf{R}}$ or $\wedge_{\mathsf{L}}$ depending on whether the variable p was in F_1 or F_2. The next step is to show that there must be rules $R_1 = \mathcal{P}_1/\Sigma_1 \Rightarrow \Pi_1$ and $R_2 = \mathcal{P}_2/\Sigma_2 \Rightarrow \Pi_2$ in $\mathcal{R}$ such that

1. $\Box p \in \Sigma_1$ and $(\Rightarrow p) \restriction_{\mathcal{C}} = \Rightarrow p$ for a restriction $\mathcal{C}$ of R_1
2. $\Box p$ in Σ_2 and $\Box q \in \Pi_2$; or $\Box q \in \Pi_2$ and $(\Box p \Rightarrow) \restriction_{\mathcal{C}} = \Box p \Rightarrow$ for a restriction $\mathcal{C}$ of R_2.

For the existence of R_1 we proceed as follows. Since the formula $\Box p \to p$ is valid in all $\mathsf{S5}$-frames, the sequent $\Box p \Rightarrow p$ must be derivable in $\mathsf{Gc}\mathcal{R}\mathsf{ConW}$ and by admissibility of Weakening therefore also derivable in $\mathsf{Gc}\mathcal{R}\mathsf{Con}$. We now consider all possible derivations of this sequent. Such a derivation must end with an application of a modal rule followed by a number of applications of Con. And since rules with restrictions introduce a layer of modalities in their principal formulae, in the last applied modal rule R the principal formulae must include $\Box p \Rightarrow$. But then for a restriction $\mathcal{C}$ of R we must have $(\Rightarrow p) \restriction_{\mathcal{C}} = \Rightarrow p$, since otherwise the sequent $\Box p \Rightarrow$ would be derivable using essentially the same derivation. The latter can not be the case, since the formula $\Box p \to \bot$ is not valid in all $\mathsf{S5}$-frames. The existence of R_2 is shown as in the proof of Theorem 2.5.6 using the $\mathsf{S5}$-valid formula $\Box p \to \Box(p \vee q)$.

Now we simplify the formulae in the context restrictions of rules in $\mathcal{R}$ in the familiar way: if e.g. for a restriction $\langle F_1, F_2 \rangle$ of a rule in $\mathcal{R}$ we have $A \vee B \in F_2$ we use mixed-cut closure of $\mathsf{Gc}\mathcal{R}$ and get that every sequent satisfying the context restriction of the rule $\vee_{\mathsf{L}}$ must also satisfy the restriction $\langle F_1, F_2 \rangle$. Thus we have $p \in F_1$ and $p \in F_2$. The other propositional connectives are analogous. If $\Box A \in F_1$, then mixed-cut closure with the rule R_2 from above gives $\Box p \in F_1$. Finally, if $\Box A \in F_2$, then mixed-cut closure with rule R_1 from above gives $p \in F_1$. Thus w.l.o.g. we may replace $\mathcal{R}$ by an equivalent set $\mathcal{R}'$ of rules where only formulae of the desired format appear in the restrictions. Obviously this preserves mixed-cut closure of the rule set. □

The translations of rules of this format are not strong enough to capture the logic $\mathsf{S5}$.

Theorem 3.4.18 (c)**.** *There is no set $\mathcal{R}$ of modal rules (not necessarily containing* Mon *or* Cong*) such that* $\mathsf{Gc}\mathcal{R}$ *is mixed-cut closed and such that* $\mathsf{Gc}\mathcal{R}\mathsf{ConW}$ *is sound and cut-free complete for* $\mathsf{S5}$*.*

Proof. Similar to Lemma 3.4.12 using the preceding Lemma 3.4.17 it is not too difficult to see that if there is such a rule set, then $\mathsf{S5}$ can be axiomatised in a Hilbert style system by translatable clauses with context formulae in $(\{\Box p, q\}, \{r\})$. The negations of such formulae are equivalent to formulae of the form

$$q \wedge \Box p \wedge P \wedge \bigwedge_{i=1}^{n} \Box A_i \wedge \bigwedge_{j=1}^{m} \neg \Box B_j \wedge \neg r$$

where for a set V of variables with $V \cap \{p, q, r\} = \emptyset$ we have $P \in \mathsf{Prop}(V)$ and for $i \leq n, j \leq m$ we have $A_i \in \mathcal{F}_r{}^{\mathsf{c}}(\{\Box p, q\}, V, \{r\})$ and $B_j \in \mathcal{F}_\ell{}^{\mathsf{c}}(\{\Box p, q\}, V, \{r\})$. Again, the formulae $q, \Box p$ and $\neg r$ need not appear in the formula, but then they may be replaced with $\top$ resp. $\Box\top$. Importantly, the formula $\Box p$ (instead of $\Box\top$) appears in the A_i or B_j only if it also appears on the top level.

Now consider the two frames $\mathfrak{F}_1 := (\mathbb{N}, \leq)$ and $\mathfrak{F}_2 := (\mathbb{N}, \mathbb{N} \times \mathbb{N})$. Suppose we have a translatable clause A with context formulae in $(\{\Box p, q\}, \{r\})$. Then $\neg A$ is equivalent to a formula of the form given above. Suppose there is a world w in $\mathfrak{F}_2$ and a valuation σ_2 on $\mathfrak{F}_2$ such that $\mathfrak{F}_2, \sigma_2, w \Vdash \neg A$. W.l.o.g. we have $w = 0$. But then defining the valuation σ_1 by setting $\sigma_1(n) := \sigma_2(n)$ for every $n \in \mathbb{N}$ we get $\mathfrak{F}_1, \sigma_1, 0 \Vdash \neg A$ as follows: Since $\mathfrak{F}_2, \sigma_2, 0 \Vdash \Box p$ we have $\mathfrak{F}_2, \sigma_2, n \Vdash p$ for every $n \in \mathbb{N}$. Thus also $\mathfrak{F}_1, \sigma_1, n \Vdash p$ for every $n \in \mathbb{N}$ and $\mathfrak{F}_1, \sigma_1, m \Vdash \Box p$ for every $m \in \mathbb{N}$. Moreover, for every $m \in \mathbb{N}$ and for every variable $s \in V \cup \{q, r\}$ we have $\mathfrak{F}_1, \sigma_1, m \Vdash s$ iff $\mathfrak{F}_2, \sigma_2, m \Vdash s$. But since every world $m \in \mathbb{N}$ is accessible from the world 0 in $\mathfrak{F}_1$ iff it is accessible from 0 in $\mathfrak{F}_2$, and since the only formula in $\neg A$ occurring under more than one modality is p, we get $\mathfrak{F}_1, \sigma_1, 0 \Vdash \neg A$ as well. The other direction is similar. Thus in total we have $\mathfrak{F}_1 \Vdash A$ iff $\mathfrak{F}_2 \Vdash A$. Now again Lemma 3.4.5 gives the result. $\Box$

Remark 3.4.19. As we have seen in Remark 2.6.6 in the presence of Contraction standard variants of sequent calculi in which the principal formulae are copied into the premisses are equivalent to sequent calculi given by rules with context restrictions as long as the underlying rules are rules with context restrictions. Thus the limitative results of this section also extend to these kinds of sequent calculi.

3.5 Notes

Axioms to rules. Our translation from axioms to rules is based on the results and methods in [Sch07, SP09], where rank-1 axioms for modal logics based on classical propositional logic are translated into one-step rules. An intermediate step, the adaption of this method to non-iterative axioms based on classical propositional logic, was published in [LP11]. Subsequently, the translation method was extended to cover non-iterative axioms for modal logics based on intuitionistic propositional logic in [LP13a], which also contains a heuristic method for the translation of nested axioms into rules with context restrictions. The full characterisation of translatable axioms for modal logics based on classical propositional logic will appear in [LP13b]. The characterisation, in particular the notion of a left- (resp. right-) resolvable formula was inspired by the definition of the substructural hierarchy in [CGT08, CGT12].

Apart from the already mentioned results for rank-1 axioms in [Sch07, SP09] there seem not to be many systematic investigations of translations from axioms into a particular format of logical sequent rules in the standard sequent setting. While e.g. translations of axioms for normal modal logics into rules of a (Hilbert-style) proof system are given in [BG13], the rule format considered there allows for arbitrary formulae occurring in the premisses and conclusion of a rule. Such rules are then turned into *reduced* rules, i.e. rules where all formulae occurring in the premisses or the conclusion are non-iterative, and where every variable occurs in the scope of a modality.

Concerning extensions of the sequent framework perhaps the best understood connection is that between modal (temporal) axioms and structural rules for extensions of modal (temporal) display logic as exhibited in [Kra96]. There we find a translation from *primitive axioms*, i.e. axioms of the form $A \to B$ where the formulae A and B contain only propositional variables, $\top, \vee, \wedge$ or the forward or backward looking temporal modalities $\Diamond_F, \Diamond_P$, and where A contains each propositional variable at most once, into structural rules such that extending the standard display calculus for temporal modal logic with these rules preserves Belnap's conditions for cut elimination. Moreover, every such structural rule ist translated back into a primitive axiom, giving an exact characterisation of those axioms which can be treated in a modal (temporal) display calculus. Also the recent work [CR13] extended methods used in the investigation of substructural logics [CGT08, CGT12] to give systematic translations for a wide class of (not necessarily modal) axioms into structural rules for a cut-free display calculus.

The line of research concerning automatic construction of cut-free labelled sequent calculi as followed e.g. in [Neg05] is based on viewing normal modal logics as fragments of first order logic. Modal axioms are first translated into corresponding frame properties expressed as geometric implications in first-order logic, then the machinery developed in [NvP98, NvP01, Neg03] is used to turn these first-order axioms into rules in the format of a *regular rule scheme*. After closing under contractions the resulting sequent calculi are guaranteed to have cut elimination. While this method is very general, it heavily relies on the correspondence of modal axioms to first-order frame conditions and on the fact that the resulting formulae are geometric implications. Thus it is not suitable e.g. for non-normal modal logics. The method has subsequently been used to construct sequent calculi for multi-modal constructive modal logics in [GGN12] and has recently been extended in [CMS13] to handle frame conditions expressed by $\forall\exists$-formulae as well.

Similarly, the recent work [Lah13] translates frame conditions for normal modal logics expressed as *n-simple* first-order formulae into hypersequent rules, which extend a basic hypersequent system for normal, transitive or symmetric frames. In the first two cases the resulting calculus has strong cut admissibility, and in the latter case it can be shown to have the subformula property. In either case the results yield decidability of the logic under scrutiny.

There also has been a fair amount of work on the problem of converting axioms for substructural logics into structural rules for a sequent or hypersequent calculus [CGT08, CGT12, CST09]. Here the axioms are decomposed using invertibility of some of the propositional rules and then turned into structural (hyper-)sequent rules. In order to identify the class of axioms which can be translated in such a way, the authors introduce the *substructural hierarchy*, which inspired our Definition 3.2.2. In case the resulting rules satisfy an additional property, that of *acyclicity*, they can be transformed into so-called *analytic* rules by a process involving restructuring of the rules to confine the active parts to the left hand sides of the sequents, and *completion* of the premisses – a process very similar to our process of variable elimination (Definition 2.4.1). Using semantical methods it is then shown that the resulting calculi enjoy a strong form of cut admissibility. These methods are extended in [CMS13] to construct a hypersequent calculus for the intermediate logic Bd_2. In contrast to the structural rules in the above mentioned works the additional rule here introduces a logical connective.

In the context of paraconsistent logics [CLSZ13] gives an automatic translation from a restricted class of Hilbert-style axioms into logical rules extending a standard sequent calculus for positive propositional logic. The axiom format is chosen in such a way that the rules resulting from the translation are amenable to the extraction of a semantics in terms of partial non-deterministic matrices introduced in [BLZ12]. This semantics ensures decidability of the logic in question and is used to check whether the resulting calculus is analytic. Furthermore, in case the resulting calculus is not analytic, the semantics is used to construct a finite family of cut-free calculi which are equivalent to the original calculus in the sense that every sequent is derivable in the latter calculus iff it is derivable in every one of the calculi in the family.

Rules to axioms. The method of using a substitution witnessing projectivity of the formula corresponding to the premisses of a rule to translate this rule into an equivalent Hilbert-style axioms was used in [Sch07] to show that one-step rules are equivalent to rank-1 axioms. Our extension of this method to cover the translation of rules with restrictions based on classical propositional logic into axioms will be published in [LP13b].

There seem not to be too many investigations into explicit translations of sequent rules into Hilbert-style axioms. One example is the procedure given in [Kra96] to translate structural rules in a modal temporal display calculus satisfying Belnap's conditions for cut elimination into primitive axioms. The idea here is to first transfer all the active parts to one side of the sequents using structural equivalences and then turn the resulting rule into an axiom using Ackemann's Lemma and the standard translation of the structural connectives into logical connectives.

Limitative results. Limitative results for particular logics and particular formats of sequent rules in the literature seem to be sparse. One example for this is the result presented in [Tiu06] stating that deep inference is necessary to capture the logic BV, an extension of multiplicative linear logic. Another example are the results in [BCar] giving sufficient conditions for when a logic can not have an $\exists$-*analytic* calculus. The definition of an $\exists$-analytic calculus is very broad and encompasses amongst cut-free sequent calculi also cut-free calculi in the framework of display logic. Thus the method probably can not be adapted to show limitative results e.g. about the logics $\mathsf{S5}$ or GL, for which such systems do exist.

4 Construction of Cut-free Sequent Systems

Given a saturated and tractable set of rules with restrictions we can use the results from Chapter 2 to show cut elimination and decidability. But in practice we are rarely given such a saturated and tractable set of rules. For example if we start with a modal logic given as a finite set of axioms for a Hilbert-style system and use the methods of Chapter 3 to convert these into sequent rules, the resulting rule set in general is not saturated. Following Question 1.1.3 from the introduction this of course gives rise to the question how to construct saturated rule sets starting from a (finite) set of rules, and whether it is possible to automate this process.

Following the idea of [PS08, PS09, PS10, PS11] the first step here is to make the rule set principal-cut closed, that is to absorb cuts on principal formulae into the rule set, by adding all possible cuts between rules. We will see in the next section how this can be done in general by considering an appropriate representation of the resulting rules. Moreover, a closer examination of the representations shows that the rules are even tractable in the sense of Definition 2.7.3. Unfortunately the resulting rule sets need not be contraction closed and mixed- and context-cut closed, and thus we do not always obtain cut elimination and decidability. Nevertheless, by restricting the format of the rules we can show some positive results, e.g. that it is always possible to construct a cut-free sequent system of our specific format for logics axiomatised by non-iterative Horn clauses. We illustrate these techniques by applications to Elgesem's logic of agency and ability as well as to weak conditional logics.

Constructing principal-cut closed rule sets by hand using this method can be a daunting prospect: we need to compute many cuts between rules, and for this we need to compute many cuts between many premisses. Since this can be very tedious and moreover is prone to errors we introduce in Section 4.2 a graphical tool to manipulate rules with restrictions and to compute cuts between rules and contractions of rules. Not only can this tool be used to manually construct cuts between rules, it also helps to spot patterns in the resulting rule set and can be used to show that a rule set is contraction closed.

4.1 Principal-cut Closure via Cut Trees

When we are faced with the task of constructing a saturated rule set from a given set of rules, the first step is to absorb cuts between principal formulae into the rule set. This idea probably was used many times in the construction of cut-free sequent systems, as suggested e.g. by the rule set $\mathcal{R}_{\mathsf{K}}$ for the standard modal logic K and countless others. As a general method it was put forward e.g. in [PS08, PS09, PS10, PS11]. With the notion of a cut between rules and with Lemma 2.4.5 we now have tools at hand, which enable us to absorb cuts into the rule set in a very systematic and purely syntactical way: since for rule sets including the propositional rules by Lemma 2.4.5 cuts between rules are derivable rules, we may simply saturate the rule set under the addition of cuts between rules. The resulting rule set then will be principal-cut closed. Of course there is a price to be paid for this: the resulting rule set in general will consist of infinitely many rules. This is not necessarily a problem if we have a tractable representation of the rules, though. For example the set $\mathcal{R}_{\mathsf{K}}$ of rules for the standard modal logic K is an infinite set of rules represented in a finite and tractable way. In this section we will introduce the concept of a *cut tree* as a possibility to represent the rules resulting from cutting a number of rules. We will see that under modest restrictions on the rule set (which can be satisfied by performing a preprocessing step) cut trees of a small size suffice to represent all possible cuts between rules from the rule set. This will give us for every finite set of rules a tractable and principal-cut closed rule set equivalent to it. Unfortunately the resulting rule sets do not necessarily satisfy the other criteria used to ensure cut elimination and generic decidability, and it is not clear whether there is a generic way to force these criteria. But if they are satisfied we automatically obtain cut elimination and decidability results. Let us try to make this more precise.

Definition 4.1.1. Let $\mathcal{R}$ be set of rules with restrictions. Define the *cut closure of* $\mathcal{R}$ to be the minimal (with respect to $\subseteq$) set $\mathsf{cc}(\mathcal{R})$ of rules with restrictions such that $\mathcal{R} \subseteq \mathsf{cc}(\mathcal{R})$ and such that for every two rules R_1, R_2 from $\mathsf{cc}(\mathcal{R})$ we have $\mathsf{cut}(R_1, R_2, \heartsuit\vec{p}) \in \mathsf{cc}(\mathcal{R})$.

Clearly the set $\mathsf{cc}(\mathcal{R})$ is principal-cut closed and can be constructed by successively adding rules $\mathsf{cut}(R_1, R_2, A)$ for rules R_1, R_1 from $\mathcal{R}$ or already constructed.

Example 4.1.2. The rule sets for standard modal logics given in Example 2.4.9 can be seen to be the cut closures of the rule sets given by translating the corresponding axioms:

1. The rule set $\mathcal{R}_{\mathsf{K}} = \{\mathsf{K}_n \mid n \geq 0\}$ is the cut closure of the rule set $\{\mathsf{K}_0, \mathsf{K}_1, \mathsf{K}_2\}$.
2. The rule set $\mathcal{R}_{\mathsf{K4}} = \mathcal{R}_{\mathsf{K}} \cup \{\mathsf{4}_n \mid n \geq 0\}$ is the cut closure of the rule set $\mathcal{R}_{\mathsf{K}} \cup \{R_4\}$.
3. The rule set $\mathcal{R}_{\mathsf{KT}} = \mathcal{R}_{\mathsf{K}} \cup \{\mathsf{T}_n \mid n \geq 1\}$ is the cut closure of $\mathcal{R}_{\mathsf{K}} \cup \{R_{\mathsf{T}}\}$.
4. The rule set $\mathcal{R}_{\mathsf{S4}} = \mathcal{R}_{\mathsf{K4}} \cup \mathcal{R}_{\mathsf{KT}}$ is the cut closure of $\mathcal{R}_{\mathsf{K}} \cup \{R_4, R_{\mathsf{T}}\}$.

In general the cut closure of a rule set will contain infinitely many rules, so we would like to represent the rules from $\mathsf{cc}(\mathcal{R})$ in a different way. We do this by making their construction from rules in $\mathcal{R}$ explicit.

Definition 4.1.3. Let $\mathcal{R}$ be a set of rules and let $\Sigma \Rightarrow \Pi$ be a sequent such that for all formulae $A, B \in \Sigma, \Pi$ we have $\mathsf{var}(A) \cap \mathsf{var}(B) = \emptyset$. An *$\mathcal{R}$-cut tree* with *principal formulae* $\Sigma \Rightarrow \Pi$ and *leafs* $R_1, \dots, R_n$ for $R_1, \dots, R_n \in \mathcal{R}$ is a proof of $\Sigma \Rightarrow \Pi$ from the principal formulae of the R_i using only Cut such that no two cut formulae share a variable. The number of nodes in a cut tree $\mathcal{D}$ is denoted by $\mathsf{size}(\mathcal{D})$, its depth by $\mathsf{depth}(\mathcal{D})$.

Note that in the definition of a cut tree we only allow applications of the cut rule, that the cut formulae all are principal formulae of the rules at the leafs of the cut tree and that no formula occurs twice as a cut formula. The restriction on the variables in the principal formulae is necessary for turning cut trees into rules.

Example 4.1.4. For $\mathcal{R} := \{\mathsf{K}_2, \mathsf{K}_3, R_{\mathsf{T}}\}$ the following is an $\mathcal{R}$-cut tree with principal formulae $\Box r_1, \Box r_2, \Box q_1, \Box q_2 \Rightarrow$ and leafs $\mathsf{K}_2, \mathsf{K}_3, R_{\mathsf{T}}$:

$$\dfrac{\dfrac{\dfrac{(r_1, r_2 \Rightarrow s; \mathcal{C}_\emptyset)}{\Box r_1, \Box r_2 \Rightarrow \Box s}\,\mathsf{K}_2 \qquad \dfrac{(s, q_1, q_2 \Rightarrow t; \mathcal{C}_\emptyset)}{\Box s, \Box q_1, \Box q_2 \Rightarrow \Box t}\,\mathsf{K}_3}{\Box r_1, \Box r_2, \Box q_1, \Box q_2 \Rightarrow \Box t}\,\mathsf{Cut} \qquad \dfrac{(t \Rightarrow\ ; \mathcal{C}_{\mathsf{id}})}{\Box t \Rightarrow}\,R_{\mathsf{T}}}{\Box r_1, \Box r_2, \Box q_1, \Box q_2 \Rightarrow}\,\mathsf{Cut}$$

where slightly abusing notation we add the rules at the leafs of the cut tree to the leafs of the derivation.

Intuitively, a cut tree can be seen as (almost) a rule by taking the principal formulae of the cut tree to be the principal formulae of the rule and by taking the premisses of the rules forming the leafs of the cut tree as its premisses. Unfortunately these premisses contain variables which do not occur in the principal formulae. But this can be taken care of using the technique of variable elimination from Section 2.4. This yields the following notion.

Definition 4.1.5. Let $\mathcal{R}$ be a set of rules and let $\mathcal{D}$ be an $\mathcal{R}$-cut tree. The rule $r(\mathcal{D})$ *represented by* $\mathcal{D}$ is defined inductively as follows. If $\mathsf{depth}(\mathcal{D}) = 0$, then $r(\mathcal{D})$ is the leaf of $\mathcal{D}$. If $\mathsf{depth}(\mathcal{D}) > 0$, then for two $\mathcal{R}$-cut trees $\mathcal{D}_1$ and $\mathcal{D}_2$ the cut tree $\mathcal{D}$ is of the form $\dfrac{\mathcal{D}_1 \quad \mathcal{D}_2}{\Sigma \Rightarrow \Pi}\,\mathsf{Cut}$, where $\Sigma \Rightarrow \Pi$ arises from the principal formulae of $\mathcal{D}_1$ and $\mathcal{D}_2$ by a cut on a formula A. Then the rule $r(\mathcal{D})$ is defined as the rule $\mathsf{cut}(r(\mathcal{D}_1), r(\mathcal{D}_2), A)$.

Example 4.1.6. For $\mathcal{R} := \{\mathsf{K}_2, \mathsf{K}_3, R_{\mathsf{T}}\}$ let $\mathcal{D}$ be the $\mathcal{R}$-cut tree

$$\dfrac{\dfrac{(r_1, r_2 \Rightarrow s; \mathcal{C}_\emptyset)}{\Box r_1, \Box r_2 \Rightarrow \Box s}\,\mathsf{K}_2 \qquad \dfrac{(s, q_1, q_2 \Rightarrow t; \mathcal{C}_\emptyset)}{\Box s, \Box q_1, \Box q_2 \Rightarrow \Box t}\,\mathsf{K}_3}{\Box r_1, \Box r_2, \Box q_1, \Box q_2 \Rightarrow \Box t}\,\mathsf{Cut}$$

Then the rule represented by this cut tree is

$$\begin{aligned} r(\mathcal{D}) &= \mathsf{cut}(\mathsf{K}_2, \mathsf{K}_3, \Box s) = \{(r_1, r_2, q_1, q_2 \Rightarrow t; \mathcal{C}_\emptyset)\}/\Box q_1, \Box q_2, \Box r_1, \Box r_2 \Rightarrow \Box t \\ &= \mathsf{K}_4\ . \end{aligned}$$

Taking $\mathcal{E}$ to be the $\mathcal{R}$-cut tree from Example 4.1.4, the rule represented by $\mathcal{E}$ thus is

$$r(\mathcal{E}) = \mathsf{cut}(r(\mathcal{D}), R_{\mathsf{T}}, \Box t) = \{(r_1, r_2, q_1, q_2 \Rightarrow ; \mathcal{C}_{\mathsf{id}})\}/\Box r_1, \Box r_2, \Box q_1, \Box q_2 \Rightarrow$$
$$= \mathsf{T}_4 \ .$$

Intuitively we construct the rule represented by a cut tree by cutting the premisses in the order given by the cuts of the cut tree. The equivalence of cut trees and cut closure is now immediate.

Lemma 4.1.7. *Let $\mathcal{R}$ be a set of rules. Then for every rule R we have that $R \in \mathsf{cc}(\mathcal{R})$ if and only if R is represented by an $\mathcal{R}$-cut tree.*

Proof. By induction on the construction of rules in $\mathsf{cc}(\mathcal{R})$ resp. $\mathcal{R}$-cut trees. □

While cut trees thus give us representations of the rules in the cut closure of a rule set, they are still not entirely what we are looking for. The problem is that the cut trees representing a rule in general might be arbitrarily big and thus might not be suitable candidates for the encodings of rules if we want to show tractability of the cut closure of a rule set. To show that only relatively small cut trees suffice to represent every rule in the cut closure of a rule set we first note that in a cut tree it is possible to change the order of the cuts without essentially changing the represented rule. Recall from Definition 2.4.6 that a rule R_1 *subsumes* a rule R_2 if the two rules have the same principal formulae and if from the premisses of an application of R_2 we can derive the premisses of the corresponding application of R_1 using only Weakening and Contraction.

Lemma 4.1.8. *Let $\mathcal{R}$ be a set of rules, and let $\mathcal{D}_1$ and $\mathcal{D}_2$ be $\mathcal{R}$-cut trees with the same principal formulae $\Sigma \Rightarrow \Pi$ and the same leafs $R_1, \dots, R_n$. Then the rule $r(\mathcal{D}_1)$ subsumes the rule $r(\mathcal{D}_2)$ and vice versa.*

Proof. Due to the restrictions on the cut formulae in a cut tree the two cut trees $\mathcal{D}_1$ and $\mathcal{D}_2$ differ only in the order of the cuts. But then by Lemma 2.4.4 from the premisses of and application of $r(\mathcal{D}_2)$ we can derive all the premisses of the corresponding application of $r(\mathcal{D}_1)$ using only Weakening and Contraction and vice versa. Thus each rule subsumes the other. □

This allows us to rearrange the cuts in a cut tree without essentially changing the represented rule. The reason why we may in general have arbitrarily large cut trees representing a rule is that if one of the premisses of an application of Cut consists of at most two formulae, then the other premiss contains at least as many formulae as the conclusion of the cut. Thus the number of formulae might increase as we move upwards along a branch in a cut tree. This problem disappears if the rule set is closed under cuts with such problematic rules.

Definition 4.1.9. A rule with restrictions is *small* if its principal formulae contain at most two formulae. A set $\mathcal{R}$ is *small-cut closed* if for every two rules R_1, R_2 from $\mathcal{R}$ such that at least one of R_1, R_2 is small we have $\mathsf{cut}(R_1, R_2, \heartsuit\vec{p}) \in \mathcal{R}$. The *small-cut closure of* $\mathcal{R}$ is the minimal (with respect to $\subseteq$) small-cut closed set $\mathsf{scc}\,(\mathcal{R})$ of rules such that $\mathcal{R} \subseteq \mathsf{scc}\,(\mathcal{R})$.

Again it is clear that the small-cut closure of a rule set can be constructed by successively adding all the missing rules. In the case of small-cut closure, however, starting from a finite set of rules we only need to add *finitely* many new rules, which can be done in a preprocessing step.

Example 4.1.10. The rule set $\mathcal{R} = \{\mathsf{K}_2, \mathsf{K}_3, R_\mathsf{T}\}$ from Examples 4.1.4 and 4.1.6 is not small-cut closed, since the rule R_T is small, but the rule $\mathsf{cut}(\mathsf{K}_2, R_\mathsf{T}, \Box t) = \{(q_1, q_2 \Rightarrow ; \mathcal{C}_\emptyset)\}/\Box q_1, \Box q_2 \Rightarrow = \mathsf{T}_2$ is not subsumed by any rule in $\mathcal{R}$. The small-cut closure of this rule set is the set

$$\begin{aligned} \mathsf{scc}\,(\mathcal{R}) &= \mathcal{R} \cup \{\mathsf{cut}(\mathsf{K}_2, R_\mathsf{T}, \Box t), \mathsf{cut}(\mathsf{K}_3, R_\mathsf{T}, \Box t)\} \\ &= \mathcal{R} \cup \{\mathsf{T}_2, \mathsf{T}_3\} \end{aligned}$$

which by construction is small-cut closed.

For small-cut closed rule sets we may simply permute problematic cuts in a cut tree up to the leafs using Lemma 4.1.8 and then replace them using the rules guaranteed by small-cut closure. This gives a cut tree without the problematic cuts. But this also means that, provided the cut tree itself has enough principal formulae, none of the rules at the leafs of the cut tree has principal formulae containing less than three formulae. Remember from Definition 2.1.8 that the *size* of a sequent $\Gamma \Rightarrow \Delta$ is the number $|\Gamma \Rightarrow \Delta| = \sum_{A \in \Gamma} \Gamma(A) + \sum_{A \in \Delta} \Delta(A)$ of formulae occurring in it counting multiplicities.

Lemma 4.1.11. *Let $\mathcal{R}$ be a set of rules. Then for every $\mathsf{scc}\,(\mathcal{R})$-cut tree $\mathcal{D}$ with principal formulae $\Sigma \Rightarrow \Pi$ there exists an $\mathsf{scc}\,(\mathcal{R})$-cut tree $\mathcal{E}$ with principal formulae $\Sigma \Rightarrow \Pi$ and leafs $R_1, \ldots, R_n$ such that*

1. *if $|\Sigma \Rightarrow \Pi| \leq 2$, then $\mathsf{depth}\,(\mathcal{E}) = 0$ (and thus $\mathcal{E}$ consists of a single leaf R_1 only)*
2. *if $|\Sigma \Rightarrow \Pi| > 2$, then the principal formulae of each of the rules $R_1, \ldots, R_n$ contain at least 3 formulae*
3. *$\mathsf{size}\,(\mathcal{E}) \leq \mathsf{size}\,(\mathcal{D})$ and the rules represented by $\mathcal{D}$ and $\mathcal{E}$ subsume each other.*

Proof. By induction on the size of the cut tree. The base case is trivial. Suppose the size of $\mathcal{D}$ is $n+1$, and that the premises of the lowermost cut are $\Gamma_1 \Rightarrow \Delta_1, A$ and $A, \Gamma_2 \Rightarrow \Delta_2$. Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be the induced $\mathsf{scc}\,(\mathcal{R})$-cut trees with principal formulae $\Gamma_1 \Rightarrow \Delta_1, A$ resp. $A, \Gamma_2 \Rightarrow \Delta_2$. Using the induction hypothesis we obtain $\mathsf{scc}\,(\mathcal{R})$-cut trees $\mathcal{E}_1$ and $\mathcal{E}_2$ with principal formulae $\Gamma_1 \Rightarrow \Delta_1, A$ and $A, \Gamma_2 \Rightarrow \Delta_2$, which have the properties mentioned in the Lemma. In particular for $i = 1, 2$ we have $\mathsf{size}\,(\mathcal{E}_i) \leq \mathsf{size}\,(\mathcal{D}_i)$. If the principal formulae of both

$r(\mathcal{E}_1)$ and $r(\mathcal{E}_2)$ contain at least three formulae each, then we are done. Otherwise, assume that the principal formulae of $r(\mathcal{E}_1)$ contain at most two formulae. Then $\mathcal{E}_1$ consists only of a leaf R_1 with principal formulae $\Gamma_1 \Rightarrow \Delta_1, A$ and the formula A occurs in the principal formulae of exactly one leaf R_2 of $\mathcal{E}_2$. By Lemma 4.1.8 we may permute the cut on A up to the leaf R_2 in $\mathcal{E}_2$, resulting in a cut tree representing a rule equivalent to $r(\mathcal{D})$. Since $\mathcal{R}$ is small-cut closed we may now replace the two leafs R_1 and R_2 and the cut on A with the leaf $\mathsf{cut}(R_1, R_2, A)$. Using the methods of Lemma 4.1.8 it is not too difficult to see that the rule represented by this cut tree is still equivalent to the original rule. The constructed cut tree might not yet have the properties specified in the Lemma, since the principal formulae of the leaf $\mathsf{cut}(R_1, R_2, A)$ might consist of less than three formulae, but since the size of the constructed cut tree is now smaller than the size of the original cut tree $\mathcal{D}$, we may simply apply the induction hypothesis again to obtain an $\mathsf{scc}\,(\mathcal{R})$-cut tree $\mathcal{E}$ with the desired properties. The remaining case that the principal formulae of $r(\mathcal{E}_2)$ contain at most two formulae is analogous. $\Box$

Example 4.1.12. Take $\mathcal{R}$ to be the rule set from the previous examples and $\mathsf{scc}\,(\mathcal{R})$ as given in Example 4.1.10. In the $\mathcal{R}$-cut tree $\mathcal{E}$ from Example 4.1.4 we may permute the problematic cut on the formula $\Box t$ up to the leafs. This yields the cut tree $\mathcal{F}$ given by

$$\dfrac{\dfrac{(r_1, r_2 \Rightarrow s; \mathcal{C}_\emptyset)}{\Box r_1, \Box r_2 \Rightarrow \Box s}\,\mathsf{K}_2 \qquad \dfrac{\dfrac{(s, q_1, q_2 \Rightarrow t; \mathcal{C}_\emptyset)}{\Box s, \Box q_1, \Box q_2 \Rightarrow \Box t}\,\mathsf{K}_3 \qquad \dfrac{(t \Rightarrow ; \mathcal{C}_{\mathsf{id}})}{\Box t \Rightarrow}\,R_{\mathsf{T}}}{\Box s, \Box q_1, \Box q_2 \Rightarrow \Box t}\,\mathsf{Cut}}{\Box r_1, \Box r_2, \Box q_1, \Box q_2 \Rightarrow}\,\mathsf{Cut}$$

Lemma 4.1.8 ensures that the rules $r(\mathcal{E})$ and $r(\mathcal{F})$ are equivalent. Now using small-cut closure of $\mathsf{scc}\,(\mathcal{R})$ this is turned into the $\mathsf{scc}\,(\mathcal{R})$-cut tree

$$\dfrac{\dfrac{(r_1, r_2 \Rightarrow s; \mathcal{C}_\emptyset)}{\Box r_1, \Box r_2 \Rightarrow \Box s}\,\mathsf{K}_2 \qquad \dfrac{(s, q_1, q_2 \Rightarrow ; \mathcal{C}_{\mathsf{id}})}{\Box s, \Box q_1, \Box q_2 \Rightarrow \Box t}\,\mathsf{cut}(\mathsf{K}_3, R_{\mathsf{T}}, \Box t)}{\Box r_1, \Box r_2, \Box q_1, \Box q_2 \Rightarrow}\,\mathsf{Cut}$$

in which there are no more problematic cuts.

In the cut trees resulting from this procedure there are no cuts where a premiss contains at most two formulae, and thus the size of the sequents always decreases when moving upwards towards the leafs on a branch in the cut tree. This allows us to bound the size of the cut tree in terms of the size of its principal formulae.

Corollary 4.1.13. *Let $\mathcal{R}$ be a set of rules and let $\Sigma \Rightarrow \Pi$ be a sequent with $|\Sigma \Rightarrow \Pi| \geq 3$. Then every rule in $\mathsf{cc}\,(\mathcal{R})$ with principal formulae $\Sigma \Rightarrow \Pi$ is equivalent to a rule represented by an $\mathsf{scc}\,(\mathcal{R})$-cut tree of size $\leq 2 \cdot |\Sigma \Rightarrow \Pi| - 5$.*

Proof. By Lemma 4.1.7 every rule R in $\mathsf{cc}\,(\mathcal{R})$ is represented by an $\mathcal{R}$-cut tree, and thus by Lemma 4.1.11 there is an equivalent rule which is represented by an $\mathcal{R}$-cut tree with the

$$\begin{tikzcd} \dfrac{(r_1, r_2 \Rightarrow s; \mathcal{C}_\emptyset)}{\Box r_1, \Box r_2 \Rightarrow \Box s}\,\mathsf{K}_2 \arrow[r, dash, "\Box s"] & \dfrac{(s, q_1, q_2 \Rightarrow t; \mathcal{C}_\emptyset)}{\Box s, \Box q_1, \Box q_2 \Rightarrow \Box t}\,\mathsf{K}_3 \arrow[r, dash, "\Box t"] & \dfrac{(t \Rightarrow\ ; \mathcal{C}_{\mathsf{id}})}{\Box t \Rightarrow}\,R_{\mathsf{T}} \end{tikzcd}$$

Figure 4.1: The cut graph of the cut trees given in Examples 4.1.4 and 4.1.12.

properties specified in the lemma. In particular, if the principal formulae of R consist of at least three formulae, then the leafs have principal formulae consisting of at least 3 formulae each. An induction on the structure of such cut trees now shows the bound on the size. $\square$

While this bound on the size of the cut trees representing a rule with given principal formulae already gives us a nice characterisation of the rules in the cut closure of a small-cut closed rule set, ultimately we would like to use the generic results from Section 2.7 to show decidability and complexity results for such rule sets. But for this we need the rule sets to be tractable, in particular rule applications with a given conclusion must have small codes, and we must be able to check reasonably fast whether a sequent is a premiss of such an application given by a code. The first requirement is already met if we take the codes of rules to be the small cut trees guaranteed by the previous corollary. Yet it is not entirely clear that we can recognise the premisses of such a rule fast enough. In order to show that this is possible we also bound the depth of the cut trees. For this we represent the cut trees in a slightly different way.

Definition 4.1.14. Let $\mathcal{R}$ be a rule set and let $\mathcal{D}$ be an $\mathcal{R}$-cut tree with principal formulae $\Sigma \Rightarrow \Pi$ and leafs $R_1, \ldots, R_n$. The *cut graph for* $\mathcal{D}$ is the graph whose nodes are the leafs of $\mathcal{D}$, and where two nodes R_i, R_j are connected by an edge if and only if a formula occurs in the principal formulae of both R_i and R_j.

Now a formula occurs in the principal formulae of two different leafs of a cut tree if and only if it is the cut formula for a cut in the cut tree. Hence two nodes in the cut graph for a cut tree are connected by an edge if and only if in the cut tree there is a cut on a formula occurring in the corresponding leafs of the cut tree. Thus cut graphs capture the "essence" of cut trees: the structure of the cuts modulo rearrangement.

Example 4.1.15. The cut graph for the cut tree $\mathcal{E}$ given in Example 4.1.4 is the graph given in Figure 4.1. Here the edges are labelled with the corresponding cut formulae. The cut tree $\mathcal{E}$ can be constructed from the graph by taking the cut corresponding to the edge labelled with the formula $\Box t$ as the lowermost cut. If instead we take the cut corresponding to the edge labelled with $\Box s$ as the lowermost cut we obtain the cut tree $\mathcal{D}$ given in Example 4.1.12.

Since two disjoint generated subtrees of a cut tree can only be joined by at most one cut, cut graphs do not contain cycles and thus can be taken to be (undirected and rooted) trees if we take an arbitrary node as the root (see e.g. [Die06] for more details on these notions). Using the following adaption of the 2-3-Lemma from [ISH65] we can show that in a cut graph we can always find an edge which divides the cut graph into two graphs of (very) roughly the same size. Here for a rooted tree $\mathcal{T}$ and a node x in $\mathcal{T}$ we denote the subtree of $\mathcal{T}$ generated by the set of nodes y for which x lies on every path from the root to y and which has x as the root by $\mathcal{T}_x$, and the number of nodes in $\mathcal{T}$ by $|\mathcal{T}|$. The *children* of a node x in a rooted tree are the nodes y such that x is the immediate predecessor of y on every path from the root to y.

Lemma 4.1.16. *Let $k \in \mathbb{N}$ and $\mathcal{T}$ be a rooted tree, such that $k+1 < |\mathcal{T}|$ and each node has at most k children. Then there is a node x in $\mathcal{T}$, such that $\left\lceil \frac{1}{k+2} \cdot |\mathcal{T}| \right\rceil \leq |\mathcal{T}_x| \leq \left\lfloor \frac{k+1}{k+2} \cdot |\mathcal{T}| \right\rfloor$.*

Proof. We construct a series $(x_0, x_1, \dots, x_d)$ of nodes in $\mathcal{T}$, such that x_0 is the root, and x_d is a leaf in the following way. Let x_0 be the root. For $i \geq 0$ and x_i not a leaf let x_{i+1} be a child of x_i, such that $|\mathcal{T}_{x_{i+1}}|$ is maximal. Since x_i has at most k children we have $|\mathcal{T}_{x_i}| \leq k \cdot |\mathcal{T}_{x_{i+1}}| + 1$. Now let

$$i_0 := \min \left\{ i \in \{0, \dots, d\} \mid |\mathcal{T}_{x_i}| < \frac{1}{k+2} \cdot |\mathcal{T}| \right\}.$$

Since $\mathcal{T}_{x_0} = \mathcal{T}$ clearly we have $i_0 > 0$. Then x_{i_0-1} is the desired node. Indeed we have

$$|\mathcal{T}_{x_{i_0-1}}| \leq k \cdot |\mathcal{T}_{x_{i_0}}| + 1 < k \cdot \frac{1}{k+2}|\mathcal{T}| + 1 \leq \frac{k}{k+2}|\mathcal{T}| + \frac{1}{k+2}|\mathcal{T}|$$

which yields the upper bound. The lower bound follows by minimality of i_0. $\square$

Since an edge in a cut graph corresponds to a cut in the underlying cut tree, this enables us to rearrange the cuts in the latter and so find an equivalent balanced cut tree with bounded depth.

Lemma 4.1.17. *Let $\mathcal{R}$ be a set of rules such that the principal formulae of every rule in $\mathcal{R}$ contain at most k formulae, and let $\Sigma \Rightarrow \Pi$ be a sequent with $|\Sigma \Rightarrow \Pi| \geq 3$. Then every rule in $\mathsf{cc}(\mathcal{R})$ with principal formulae $\Sigma \Rightarrow \Pi$ is subsumed by a rule represented by an $\mathsf{scc}(\mathcal{R})$-cut tree of size at most $2 \cdot |\Sigma \Rightarrow \Pi| - 5$ and depth at most $c_k \cdot \log_2 |\Sigma \Rightarrow \Pi| + k$ for $c_k := (\log_2 \frac{k+2}{k+1})^{-1}$.*

Proof. Let R be a rule in $\mathcal{R}$. We show how to construct a cut tree $\mathcal{E}$ with the desired properties. By Corollary 4.1.13 we know that the rule R is equivalent to a rule represented by an $\mathsf{scc}(\mathcal{R})$-cut tree $\mathcal{D}$ with principal formulae $\Sigma \Rightarrow \Pi$ and size at most $2 \cdot |\Sigma \Rightarrow \Pi| - 5$. Thus the cut tree $\mathcal{D}$ has at most $|\Sigma \Rightarrow \Pi| - 2$ leafs, and the cut graph $\mathcal{G}$ for $\mathcal{D}$ has at most that many nodes. Furthermore, since the principal formulae of every rule in $\mathcal{R}$ contain at most k formulae and since cuts with small rules do not increase the number of principal formulae of a rule, the principal formulae of every rule in $\mathsf{scc}(\mathcal{R})$ contain at most k formulae as well. Thus, taking an arbitrary node in $\mathcal{G}$ as the root, every node in $\mathcal{G}$ has at most k children. This means that we may apply Lemma 4.1.16 to find a node x of $\mathcal{G}$ such that

$$\left\lceil \frac{1}{k+2} \cdot |\mathcal{G}| \right\rceil \leq |\mathcal{G}_x| \leq \left\lfloor \frac{k+1}{k+2} \cdot |\mathcal{G}| \right\rfloor .$$

Then writing $\mathcal{G} \smallsetminus \mathcal{G}_x$ for the rooted subtree of $\mathcal{G}$ generated by the complement of the set of nodes of $\mathcal{G}_x$ in G we also have $|\mathcal{G} \smallsetminus \mathcal{G}_x| \leq \left\lfloor \frac{k+1}{k+2} \cdot |\mathcal{G}| \right\rfloor$. The new cut tree $\mathcal{E}$ (with principal formulae $\Sigma \Rightarrow \Pi$) is now constructed by taking as the lowermost cut the cut in $\mathcal{D}$ corresponding to the (unique) edge connecting $\mathcal{G}_x$ and $\mathcal{G} \smallsetminus \mathcal{G}_x$. For the next level of cuts we repeat the process for the cut graphs $\mathcal{G}_x$ and $\mathcal{G} \smallsetminus \mathcal{G}_x$. Continuing upwards in this fashion after at most $\log_{\frac{k+2}{k+1}}(|\Sigma \Rightarrow \Pi| - 2)$ steps we arrive at a sub-graph with at most $k+1$ nodes, which we insert into $\mathcal{E}$ in the form of a sub-derivation of depth at most k. Thus the cut tree $\mathcal{E}$ has depth at most

$$\log_{\frac{k+2}{k+1}}(|\Sigma \Rightarrow \Pi| - 2) + k \leq \left(\log_2 \frac{k+2}{k+1}\right)^{-1} \cdot \log_2 |\Sigma \Rightarrow \Pi| + k \ .$$

By construction the size of $\mathcal{E}$ is the same as the size of $\mathcal{D}$, and by Lemma 4.1.8 the represented rules are equivalent. ☐

Thus for finite small-cut closed rule sets we have small representations of all the rules in their cut closure. But we even get a bit more. If we start with a finite set of rules, then computing the small-cut closure of this rule set only adds finitely many new rules. Moreover, this can be done in a preprocessing step, and thus only adds a constant time overhead to the construction of the individual rules. This gives us a representation of the cut closure for *every* finite set of rules.

Definition 4.1.18. Let $\mathcal{R}$ be a set of rules whose principal formulae contain at most k formulae each. The *rule set generated by* $\mathcal{R}$ is the set $\mathcal{R}^{\clubsuit}$ of rules represented by $\mathsf{scc}\,(\mathcal{R})$-cut trees with depth at most $c_k \cdot \log_2 |\Sigma \Rightarrow \Pi| + k$ where $\Sigma \Rightarrow \Pi$ are their principal formulae and $c_k = (\log_2 \frac{k+2}{k+1})^{-1}$ is the constant introduced in Lemma 4.1.17.

Theorem 4.1.19. *Let $\mathcal{R}$ be a finite set of rules with $\mathsf{G[cim]} \subseteq \mathcal{R}$. Then the rule set $\mathcal{R}^{\clubsuit}$ is principal-cut closed, tractable and equivalent to $\mathcal{R}$.*

Proof. Since the rule set $\mathcal{R}$ contains the propositional rules, equivalence of $\mathcal{R}$ and $\mathsf{scc}\,(\mathcal{R})$ follows by Lemma 2.4.5. Since $\mathcal{R}$ is finite, we know that for some k the principal formulae of every rule in $\mathcal{R}$ contain at most k formulae. Thus by Lemma 4.1.17 we have equivalence of $\mathsf{cc}\,(\mathcal{R})$ and $\mathcal{R}^{\clubsuit}$. Equivalence of $\mathcal{R}$ and $\mathsf{cc}\,(\mathcal{R})$ follows from Lemma 2.4.5.

To see that $\mathcal{R}^{\clubsuit}$ is principal-cut closed consider two rules R_1, R_2 from $\mathcal{R}^{\clubsuit}$. Since both of these rules are represented by $\mathsf{scc}\,(\mathcal{R})$-cut trees, the rule $\mathsf{cut}(R_1, R_2, \heartsuit\vec{p})$ is also represented by a $\mathsf{scc}\,(\mathcal{R})$-cut tree, and again using Lemma 4.1.17 there is a rule represented by an $\mathsf{scc}\,(\mathcal{R})$-cut tree of bounded size and depth which subsumes the rule $\mathsf{cut}(R_1, R_2, \heartsuit\vec{p})$.

To see that the rule set is tractable we take the encodings of applications of rules in $\mathcal{R}^{\clubsuit}$ to be the small cut trees representing rules in $\mathcal{R}^{\clubsuit}$ together with a substitution of their principal formulae and a context. Then given a sequent $\Gamma \Rightarrow \Delta$ and such an encoding of an application of a rule in $\mathcal{R}^{\clubsuit}$ we can check in space polynomial in $|\Gamma \Rightarrow \Delta|$ that the conclusion of the application of the rule given by the encoding is $\Gamma \Rightarrow \Delta$. Furthermore, since the cut-trees have depth at most $c_k \cdot \log_2 |\Gamma \Rightarrow \Delta| + k$, each of the premises of the represented rules is constructed by cutting at most $\left(2^{\ell}\right)^{c_k \cdot \log_2 |\Gamma \Rightarrow \Delta| + k} = |\Gamma \Rightarrow \Delta|^{\ell \cdot c_k} \cdot 2^{\ell \cdot k}$ many premises of rule applications at the leaves, where ℓ is the maximal arity of modalities. Thus given a sequent we can check whether it is a premiss of a rule application given by a small cut tree, a substitution of its principal formulae and a context by guessing at most this number of premisses of rules at the leafs of the cut tree, performing the cuts according to the cut tree, substituting the variables and adding the context. Since the values of k, c_k and ℓ only depend on the rule set and not on the input, the check can be done in time polynomial in the size of the encoding of the rule application. Therefore we can check whether a sequent is a premiss of the application in nondeterministic polynomial time and thus in polynomial space (in the size of the encoding of the rule application). $\square$

Ideally, we would like to use this representation of the cut closure to decide derivability for logics given by finite sets of rules. Unfortunately, for this we also need the rule set to be closed under contractions and context- and mixed-cuts. These are properties which are not automatically guaranteed by the construction of the rules in $\mathcal{R}^{\clubsuit}$.

Corollary 4.1.20. *Let $\mathcal{R}$ be a finite set of rules. If the rule set $\mathcal{R}^{\clubsuit}$ is right- or left-contraction closed, and mixed- and context-cut closed, then $\mathcal{R}^{\clubsuit}\mathsf{Con}$ has cut elimination. If $\mathcal{R}^{\clubsuit}$ is contraction closed and mixed- and context-cut closed, then the derivability problem for $\mathcal{R}$ is in* EXPTIME.

Proof. Cut elimination follows from Theorems 2.4.16 and 4.1.19, and the complexity bound follows from Corollary 2.7.6. $\square$

It is not clear whether the rule set can be automatically modified to satisfy these additional conditions. Since in contrast to principal-cut closure the properties of mixed- and context-cut closure are not monotone in the sense that adding a rule might destroy these properties, it is doubtful whether we can simply saturate the rule set as in the case of principal-cut closure. For contraction closure we could simply add all the missing rules, but it is not clear how to do this in a way that the resulting rule set is still tractable. Some of these obstacles can be evaded by restricting the rule format to that of *shallow* rules. Since sets of shallow rules are always mixed- and context-cut closed (Example 2.4.12), for these rule sets we only need to require contraction closure. In case we have contraction closure, then by Theorem 2.7.8 we obtain an even lower complexity bound.

Corollary 4.1.21. *Let $\mathcal{R}$ be a finite set of shallow rules. If the rule set $\mathcal{R}^{\clubsuit}$ is contraction closed, then $\mathcal{R}^{\clubsuit}\mathsf{Con}$ has cut elimination and the derivability problem for $\mathcal{R}$ is in* PSPACE. $\square$

Finally, if the principal formulae of the rules in a rule set contain at most one formula on the right hand side, then the rule set is trivially right-contraction closed. Thus for such rule sets we automatically obtain an equivalent cut-free sequent system. We can also characterise the corresponding class of axioms for Hilbert-style systems.

Definition 4.1.22. A formula is a *Horn clause* if it is of the form $\bigwedge_{i=1}^{n} B_i \to C$ for $n \geq 0$ and formulae $B_1, \ldots, B_n, C$.

Theorem 4.1.23. *Let $\mathcal{R}$ be a finite set of shallow rules where the right hand sides of the principal formulae of each rule contain at most one formula. Then $\mathcal{R}^{\clubsuit}\mathsf{Con}$ has cut elimination. Thus if $\mathcal{A}$ is a finite set of non-iterative translatable Horn clauses and if $\mathcal{R}_{\mathcal{A}}$ consists of the translations of the axioms in $\mathcal{A}$ according to Section 3.1, then the sequent system given by $(\mathsf{G[cim]}\mathcal{R}_{\mathcal{A}})^{\clubsuit}\mathsf{Con}$ has cut elimination.*

Proof. Cuts between shallow rules with at most one formula in the right hand side of the principal formulae also are shallow rules and have at most one formula on the right hand side of the principal formulae. Thus the generated rule set $\mathcal{R}^{\clubsuit}$ is context- and mixed cut closed (compare Example 2.4.12) and right-contraction closed (compare Example 2.4.14). Hence by Theorems 2.4.16 and 4.1.19 it has cut elimination. Given a finite set of translatable non-iterative Horn clauses, the translation from Section 3.2 gives an equivalent finite set of shallow rules (compare Corollary 3.2.26). It is not hard to see that the right hand sides of the principal formulae of these rules contain at most one formula. □

Thus in particular we have cut-free sequent systems for every modal logic based on intuitionistic propositional logic and axiomatised by non-iterative translatable Horn clauses and for every modal logic based on classical logic and axiomatised by non-iterative Horn clauses.

Example 4.1.24.

1. The rule set $\{\mathsf{K}_2\}^{\clubsuit}$ is (classically and intuitionistically) equivalent to the axiom $\Box p \wedge \Box q \to \Box r$ and $\mathsf{G[ci]}\{\mathsf{K}_2\}^{\clubsuit}\mathsf{Con}$ has cut elimination.

2. We have seen in Example 3.2.17 that the axioms $\Box p \wedge \Box q \to \Box(p \wedge q)$ and $(\mathsf{IK2})$ $\Box(p \to q) \wedge \Diamond p \to \Diamond q$ for constructive modal logic CK [Wij90, BdPR01] translate into the rules

$$\mathsf{K}_2 \ \{(p_1, p_2 \Rightarrow q; \mathcal{C}_\emptyset)\}/\Box p_1, \Box p_2 \Rightarrow \Box q$$
$$R_{\mathsf{IK2}} \ \{(p, q \Rightarrow r; \mathcal{C}_\emptyset)\}/\Box p, \Diamond q \Rightarrow \Diamond r$$

   Furthermore, it is not hard to see that the axiom $\Box\top$ translates into the rule $\mathsf{K}_0 = \{(\ \Rightarrow p; \mathcal{C}_\emptyset)\}/\ \Rightarrow \Box p$. Thus setting $\mathcal{R}_{\mathsf{CK}} := \{\mathsf{K}_2, R_{\mathsf{IK2}}, \mathsf{K}_0\}$ we have that the rule set $\mathsf{Gi}\mathcal{R}^{\clubsuit}_{\mathsf{CMK}}\mathsf{Con}$ is sound and complete for constructive modal logic CMK and has cut elimination.

More examples are given by weak systems of conditional logics [Che75, Che80, Nut80, OPS07, PS11]. Here we consider the logics in a purely syntactical way. Some stronger conditional logics will be examined in more detail in Chapter 5.

Example 4.1.25 (c)**.** Let $\Lambda_{\Box\!\!\to}$ be the set of connectives containing in addition to the standard boolean connectives the binary connective $\Box\!\!\to$. We write this connective in infix notation and

(CA)	$(p \mathrel{\Box\!\!\to} r) \wedge (q \mathrel{\Box\!\!\to} r) \to (p \vee q \mathrel{\Box\!\!\to} r)$
(CC)	$(p \mathrel{\Box\!\!\to} q) \wedge (p \mathrel{\Box\!\!\to} r) \to (p \mathrel{\Box\!\!\to} q \wedge r)$
(CEM)	$(p \mathrel{\Box\!\!\to} q) \vee (p \mathrel{\Box\!\!\to} \neg q)$
(CM)	$(p \mathrel{\Box\!\!\to} (q \wedge r)) \to (p \mathrel{\Box\!\!\to} q) \wedge (p \mathrel{\Box\!\!\to} r)$
(CMon)	$(p \mathrel{\Box\!\!\to} q) \wedge (p \mathrel{\Box\!\!\to} r) \to ((p \wedge q) \mathrel{\Box\!\!\to} r)$
(CN)	$p \mathrel{\Box\!\!\to} \top$
(CS)	$p \wedge q \to (p \mathrel{\Box\!\!\to} q)$
(CSO)	$(p \mathrel{\Box\!\!\to} q) \wedge (q \mathrel{\Box\!\!\to} p) \to ((p \mathrel{\Box\!\!\to} r) \leftrightarrow (q \mathrel{\Box\!\!\to} r))$
(CV)	$(p \mathrel{\Box\!\!\to} q) \wedge \neg(p \mathrel{\Box\!\!\to} \neg r) \to ((p \wedge r) \mathrel{\Box\!\!\to} q)$
(ID)	$p \mathrel{\Box\!\!\to} p$
(MP)	$(p \mathrel{\Box\!\!\to} q) \to (p \to q)$

Table 4.1: Hilbert-style axiomatisations for conditional logics formulated using $\mathrel{\Box\!\!\to}$.

read the formula $A \mathrel{\Box\!\!\to} B$ as "A conditionally implies B". Consider the axioms for conditional logics based on classical propositional logic given in Table 4.1. Since each of the axioms in

$$S := \{(\mathsf{CA}), (\mathsf{CC}), (\mathsf{CM}), (\mathsf{CMon}), (\mathsf{CN}), (\mathsf{CS}), (\mathsf{CSO}), (\mathsf{ID}), (\mathsf{MP})\}$$

is equivalent to finitely many non-iterative translatable Horn clauses, by the methods of Chapter 3 for every set $\mathcal{A}$ of axioms with $\mathcal{A} \subseteq S$ we have a corresponding set $\mathcal{R}_{\mathcal{A}}$ of translations of these axioms into rules. Thus for such a set $\mathcal{A}$ the sequent system given by the rules $\mathcal{R}_{\mathcal{A}}^{\diamond}$ is equivalent to $\mathcal{A}$ and by Theorem 4.1.23 the sequent system given by $\mathsf{Gc}\mathcal{R}_{\mathcal{A}}^{\diamond}\mathsf{Con}$ has cut elimination.

Of course now the question is how it is possible to check whether the cut closure of a rule set is saturated. In particular, the rule set needs to be contraction closed. The following lemma allows us to restrict the check to rules of a specific form. Remember from Lemma 2.4.4 that for a set $\mathcal{P}$ of premisses and a sequent $\Theta \Rightarrow \Xi$ the set $\mathcal{P}(\Theta \Rightarrow \Xi)$ contains all sequents $\Theta \restriction_{F_1}, \Sigma \Rightarrow \Pi, \Xi \restriction_{F_2}$ for premisses $(\Sigma \Rightarrow \Pi; \langle F_1, F_2 \rangle)$ in $\mathcal{P}$.

Lemma 4.1.26. *Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be sets of premisses, and let $\Theta \Rightarrow \Xi$ be a sequent. If every sequent in $\mathcal{P}_2(\Theta \Rightarrow \Xi)$ is derivable from $\mathcal{P}_1(\Theta \Rightarrow \Xi)$ using only* ConW*, then for every variable p every sequent in $(\mathcal{P}_2 \ominus p)(\Theta \Rightarrow \Xi)$ is derivable from $(\mathcal{P}_1 \ominus p)(\Theta \Rightarrow \Xi)$ using only* ConW.

Proof. Take a sequent $\Gamma \Rightarrow \Delta$ from $(\mathcal{P}_2 \ominus p)(\Theta \Rightarrow \Xi)$. We look at the possible cases.

If $\Gamma \Rightarrow \Delta \in \mathcal{P}_2(\Theta \Rightarrow \Xi)$, then p does not occur in $\Gamma \Rightarrow \Delta$. Moreover, since the sequent is derivable from $\mathcal{P}_1(\Theta \Rightarrow \Xi)$ using only ConW there is a sequent $\Gamma' \Rightarrow \Delta'$ from $\mathcal{P}_1(\Theta \Rightarrow \Xi)$ with

$\mathsf{Supp}(\Gamma') \subseteq \mathsf{Supp}(\Gamma)$ and $\mathsf{Supp}(\Delta') \subseteq \mathsf{Supp}(\Delta)$. Thus p does not occur in $\Gamma' \Rightarrow \Delta'$ either and we have $\Gamma' \Rightarrow \Delta' \in (\mathcal{P}_1 \ominus p)(\Theta \Rightarrow \Xi)$. Thus $\Gamma \Rightarrow \Delta$ follows from $(\mathcal{P}_1 \ominus p)(\Theta \Rightarrow \Xi)$ using only ConW.

Otherwise we have $\Gamma \Rightarrow \Delta = \Sigma_1, \Sigma_2 \Rightarrow \Pi_1, \Pi_2$ where the sequents $\Sigma_1 \Rightarrow \Pi_1, p$ and $p, \Sigma_2 \Rightarrow \Pi_2$ are in $(\mathcal{P}_2 \ominus p)(\Theta \Rightarrow \Xi)$. Thus we have sequents $\Sigma_1' \Rightarrow \Pi_1'$ and $\Sigma_2' \Rightarrow \Pi_2'$ in $\mathcal{P}_1(\Theta \Rightarrow \Xi)$ with

$$\begin{array}{rl} \mathsf{Supp}\left(\Sigma_1'\right) \subseteq \mathsf{Supp}(\Sigma_1) & \quad \mathsf{Supp}\left(\Pi_1'\right) \subseteq \mathsf{Supp}(\Pi_1, p) \\ \mathsf{Supp}\left(\Sigma_2'\right) \subseteq \mathsf{Supp}(p, \Sigma_2) & \quad \mathsf{Supp}\left(\Pi_2'\right) \subseteq \mathsf{Supp}(\Pi_2)\ . \end{array}$$

But then at least one of the three sequents

$$\Sigma_1' \Rightarrow \Pi_1' \qquad \Sigma_2' \Rightarrow \Pi_2' \qquad \Sigma_1', \Sigma_2' \Rightarrow \Pi_1', \Pi_2'$$

must be in $(\mathcal{P}_1 \ominus p)(\Theta \Rightarrow \Xi)$, and thus the sequent $\Gamma \Rightarrow \Delta$ follows from $(\mathcal{P}_1 \ominus p)(\Theta \Rightarrow \Xi)$ using only ConW. □

Using this lemma we can show that it is enough to only consider contractions along a path in a cut graph.

Definition 4.1.27. Let $\mathcal{R}$ be a set of rules. A rule from $\mathsf{cc}(\mathcal{R})$ is called a *path rule*, if it is represented by a cut tree whose cut graph is a path.

Example 4.1.28. The rules represented by the cut trees given in Examples 4.1.4 and 4.1.12 are path rules, since the cut graph for these cut trees (shown in Figure 4.1) is a path.

In particular, if R is a leaf of a cut tree witnessing that a rule is a path rule, then at most two formulae of the principal formulae of R occur as cut formulae in the cut tree.

Theorem 4.1.29. *Let $\mathcal{R}$ be a set of rules. Suppose that all contractions of path rules from $\mathsf{cc}(\mathcal{R})$ are subsumed by rules in $\mathsf{cc}(\mathcal{R})$. Then $\mathsf{cc}(\mathcal{R})$ is contraction closed.*

Proof. Let R be a rule in $\mathsf{cc}(\mathcal{R})$. We need to show that every contraction of R is subsumed by a rule in $\mathsf{cc}(\mathcal{R})$. By Lemma 4.1.7 the rule R is represented by a cut tree $\mathcal{D}$. Let $\mathcal{G}$ be the cut graph for $\mathcal{D}$. Consider the (left or right) contraction of R on the literals $\heartsuit\vec{p}$ and $\heartsuit\vec{q}$. These two literals occur in the principal formulae of exactly one leaf of $\mathcal{D}$ each. Let $\mathcal{H}$ be the cut graph induced by the path connecting the two nodes of $\mathcal{G}$ corresponding to these two leafs. Then imposing an arbitrary ordering on the cuts from $\mathcal{H}$ we obtain a cut tree $\mathcal{E}$ which represents a path rule Q in $\mathsf{cc}(\mathcal{R})$. But since $\mathsf{cc}(\mathcal{R})$ is contraction closed for path rules, the contraction of Q on $\heartsuit\vec{p}$ and $\heartsuit\vec{q}$ is subsumed by a rule S in $\mathsf{cc}(\mathcal{R})$. Let $\mathcal{H}'$ be the cut graph for the cut tree for the rule S. Then since the rule S has the same principal formulae as the contraction of the rule Q, in the cut graph $\mathcal{G}$ we may replace the path $\mathcal{H}$ by $\mathcal{H}'$, connecting the remaining parts of $\mathcal{G}$ to those nodes of $\mathcal{H}'$ where the cut formula corresponding to the according edge occurs in the principal formulae. Again, taking an arbitrary ordering of the cuts we obtain a cut tree. Let T be the rule represented by this cut tree. Then the rule T has the same principal formulae as the contraction of the rule R on $\heartsuit\vec{p}$ and $\heartsuit\vec{q}$, and by repeated applications of Lemma 4.1.26 all the premisses of the latter are derivable from the premisses of T using only ConW. Thus the contraction of R is subsumed by a rule in $\mathsf{cc}\,(\mathcal{R})$. □

In particular, since rules with at most one formula in the principal formulae can not be used in the construction of the path rules used in the proof of Theorem 4.1.29, such rules can be added to a rule set without destroying contraction closure of the cut closure.

Corollary 4.1.30. *Let $\mathcal{R}$ be a set of rules such that $\mathsf{cc}\,(\mathcal{R})$ is contraction closed. Then for every rule R with principal formulae $\Sigma \Rightarrow \Pi$ such that $|\Sigma \Rightarrow \Pi| = 1$ the set $\mathsf{cc}\,(\mathcal{R}R)$ is contraction closed.*

Proof. Since the rule R has only one principal formula, nodes in a cut graph labelled with R have degree one. Moreover, in the cut tree representing a rule Q no principal formulae of Q occur in a leaf labelled with R. Thus when constructing the cut graph $\mathcal{H}$ in the proof of Theorem 4.1.29 above, no node of $\mathcal{H}$ is labelled with R. Reasoning as above, contraction closure of $\mathsf{cc}\,(\mathcal{R}R)$ is now seen to depend only on contraction closure of $\mathsf{cc}\,(\mathcal{R})$. □

This is an important result, since it means that in order to check whether the cut closure of a rule set is contraction closed it is sufficient to check contraction closure of the cut closure of the subset consisting of 'big' rules, i.e. rules with at least three formulae in the principal formulae. This also means that once we have established contraction closure of the cut closure of a rule set we may add 'very small' rules, i.e. rules whose principal formulae consist of only one formula, for free. We can use this to uniformly establish contraction closure for some basic systems of conditional logic and thus uniformly re-establish the decidability and complexity results for these systems found in [Che75, OS01, OPS07, PS09, PS11].

Theorem 4.1.31 (c)**.** *Let $\mathcal{A} \subseteq \{(\mathsf{CC}), (\mathsf{CEM}), (\mathsf{CM}), (\mathsf{CN}), (\mathsf{CS}), (\mathsf{ID}), (\mathsf{MP})\}$ with $(\mathsf{CM}) \in \mathcal{A}$ and with $(\mathsf{CN}) \in \mathcal{A}$ whenever $(\mathsf{CEM}) \in \mathcal{A}$. Then if $\mathcal{R}_\mathcal{A}$ is the set of rules consisting of the translations of the axioms in $\mathcal{A}$ the rule set $(\mathcal{R}_\mathcal{A})^{\oplus}$ is contraction closed. Moreover, the sequent calculus given by $\mathsf{Gc}(\mathcal{R}_\mathcal{A})^{\oplus}\mathsf{Con}$ has cut elimination and the derivability problem for this system is in* PSPACE.

Proof. Translating the axiom (CM) of Table 4.1 into a rule using the methods of Section 3.2 (see in particular Remark 3.2.27) gives the rule

$$R_{\mathsf{CM}} \quad \{(p \Rightarrow q; \mathcal{C}_\emptyset), (q \Rightarrow p; \mathcal{C}_\emptyset), (r \Rightarrow s; \mathcal{C}_\emptyset)\}/p \mathrel{\Box\!\!\rightarrow} r \Rightarrow q \mathrel{\Box\!\!\rightarrow} s$$

which states monotonicity of $\Box\!\!\rightarrow$ in the second argument. Using this to translate the axioms (CC), (CEM) and (CN) into rules we obtain the rules

$$
\begin{array}{rl}
R_{\mathsf{CC}} & \{(p_i \Rightarrow p_j; \mathcal{C}_\emptyset) \mid i,j \in \{1,2,3\}, i \neq j\} \cup \{(q,r \Rightarrow s; \mathcal{C}_\emptyset)\} / p_1 \Box\!\!\rightarrow q, p_2 \Box\!\!\rightarrow r \Rightarrow p_3 \Box\!\!\rightarrow s \\
R_{\mathsf{CEM}} & \{(p \Rightarrow q; \mathcal{C}_\emptyset), (q \Rightarrow p; \mathcal{C}_\emptyset), (\Rightarrow s,t; \mathcal{C}_\emptyset)\} / \Rightarrow p \Box\!\!\rightarrow s, q \Box\!\!\rightarrow t \\
R_{\mathsf{CN}} & \{(\Rightarrow q; \mathcal{C}_\emptyset)\} / \Rightarrow p \Box\!\!\rightarrow q\ .
\end{array}
$$

Note that the rule R_{CN} is the right contraction of R_{CEM}. It is now not too difficult to see that path rules (or indeed any rules) in $\mathsf{cc}(\{R_{\mathsf{CM}}, R_{\mathsf{CC}}, R_{\mathsf{CEM}}, R_{\mathsf{CN}}\})$ have the form $\mathsf{CKCEM}_{n,m} = \mathcal{P}_{n,m}/p_1 \Box\!\!\rightarrow q_1, \dots, p_n \Box\!\!\rightarrow q_n \Rightarrow p_{n+1} \Box\!\!\rightarrow q_{n+1}, \dots, p_{n+m} \Box\!\!\rightarrow q_{n+m}$ for $n \geq 0, m \geq 1$ with

$$
\mathcal{P}_{n,m} := \{(p_i \Rightarrow p_j; \mathcal{C}_\emptyset) \mid i,j \leq n+m, i \neq j\} \cup \{(q_1, \dots, q_n \Rightarrow q_{n+1}, \dots, q_{n+m}; \mathcal{C}_\emptyset)\}
$$

and that path rules (or again any rules) in $\mathsf{cc}(\{R_{\mathsf{CM}}, R_{\mathsf{CC}}\})$ have the form $\mathsf{CK_n} := \mathsf{CKCEM}_{n,1}$ for $n \geq 1$, and thus that both of these rule sets are contraction closed (see also [Che80, PS08, PS09, PS10, PS11]). Thus by Corollary 4.1.30 all extensions of these rule sets with rules whose principal formulae contain at most one formula are contraction closed as well. This gives contraction closure and cut elimination for the rule sets $\mathsf{Gc}\mathcal{R}^{\clubsuit}_{\mathcal{A}}\mathsf{Con}$ for all sets $\mathcal{A}$ as stated in the theorem. The complexity result then follows from the results in Section 2.7. $\Box$

4.2 Sequent Rules in Pictures

The results of the previous section show that it is possible to construct principal-cut closed and tractable rule sets for many modal logics in an automatic way by considering a representation of the rules in the cut closure of a rule set as cut trees. Unfortunately the construction does not guarantee contraction closure or context- or mixed-cut closure of the resulting rule set. In order to check whether e.g. contraction closure holds, it is sometimes advisable to explicitly construct all the rules in the rule set, as done e.g. in the proof of Theorem 4.1.31. Moreover, in some cases we might prefer an explicit representation of the rules to their representation in terms of cut trees, e.g. if we would like to use the rules in a very efficient decision procedure. Thus sometimes we still need to construct the principal-cut closure of a rule set by hand. One of the standard methods for this is to saturate the rule set under the addition of cuts between rules until a pattern in the form of the rules can be distinguished. As can be testified by anyone who tried to do this for a reasonably complex logic, due to the sheer number of sequents which need to be handled in each step, this can be a very laborious and error-prone process. This holds especially if the modalities have arity greater than one or are not monotone.

To address this issue we now introduce a graphical tool to aid the manual handling of rules, cuts between rules and contractions of rules. It should be noted that 'tool' here is meant in a theoretical sense and not in the sense of an implementation. The general idea is to represent sequents as *doodles*, that is as arrows with multiple heads and tails. Then performing a cut between two sequents amounts to connecting heads and tails of arrows, and applying Contraction is mirrored by identifying two heads or two tails. This presentation is extended in a natural way to rules and the operations of cut between rules and contraction of a rule. There are two major advantages to this presentation: The first is that as mentioned earlier this presentation greatly reduces the amount of effort required for computing cuts between rules. The second advantage is that it allows us to spot patterns in the construction of rules more easily. Being able to do so is very important, since the principal-cut closure of a rule set in general comprises infinitely many rules of many different shapes, and thus we need to identify a suitably regular subset of rules which subsumes all the rules in the constructed set. The examples in this section mainly serve illustrative purposes, and thus are relatively simple examples such as the standard normal modal logics or the weak conditional logics considered in the last section. The tool will be put to full use in the next chapter, where we will construct new cut-free sequent calculi for stronger conditional logics. Let us start with the basic notion.

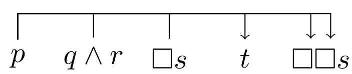
Definition 4.2.1. Let F be a set of formulae. A *doodle over* F is an arrow with multiple heads and tails pointing to (resp. emerging from) formulae in F. A formula occurs *positively* in a doodle if a head points to it, and *negatively* if a tail emerges from it. A doodle d *represents a sequent* $\Gamma \Rightarrow \Delta$ if exactly the formulae in Γ occur negatively in d and exactly the formulae in Δ occur positively in d and if moreover the multiplicity of every formula in Γ is the number of heads pointing to it in d and the multiplicity of every formula in Δ is the number of tails emerging from it in d.

It is clear that sequents and doodles are in one-to-one correspondence.

Fact 4.2.2. *Every sequent is represented by exactly one doodle and every doodle represents exactly one sequent.*

Using this simple correspondence we represent sequents in a graphical way.

Example 4.2.3. The following is a doodle over the set $\{p, q \wedge r, \Box s, t, \Box\Box s\}$ of formulae.

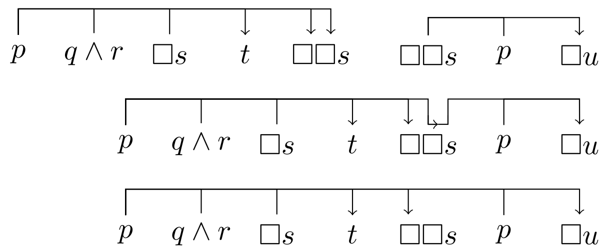


The formulae $p, q \wedge r, \Box s$ occur negatively in the doodle, the formulae t and $\Box\Box s$ positively. The sequent represented is $p, q \wedge r, \Box s \Rightarrow t, \Box\Box s, \Box\Box s$.

This allows us to formulate the graphical counterpart of a cut between sequents in an intuitive way as the result of connecting heads and tails of doodles.

Definition 4.2.4. Let A be a formula and let d_1 and d_2 be doodles such that A occurs positively in d_1 and negatively in d_2. Then the *cut between* d_1 *and* d_2 *on* A is the doodle arising from d_1 and d_2 by connecting one head of d_1 pointing to A to one tail of d_2 emerging from A.

Example 4.2.5. Let d_1 be the doodle from Example 4.2.3 and let d_2 be the doodle representing the sequent $\Box\Box s, p \Rightarrow \Box u$. Then the cut between d_1 and d_2 on $\Box\Box s$ is constructed the following way. We start with the two doodles d_1 and d_2 as given in (4.1) below.



(4.1)

(4.2)

(4.3)

Now we identify the occurrences of the cut formula $\Box\Box s$ and connect one head of d_1 pointing to this formula to one tail of d_2 emerging from it as in (4.2). Finally we 'yank the wire', which gives the doodle shown in (4.3). In this case we still have a head pointing to the formula $\Box\Box s$. If this is not the case we also omit the formula.

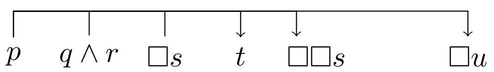
Again it is clear that performing cuts on sequents corresponds to performing cuts on doodles.

Fact 4.2.6. *If the doodles d_1 and d_2 correspond to the sequents $\Gamma_1 \Rightarrow \Delta_1, A$ and $A, \Gamma_2 \Rightarrow \Delta_2$, then the cut between d_1 and d_2 on A corresponds to the sequent $\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2$.*

Similarly, we have a graphical analogue to applying Contraction to a sequent.

Definition 4.2.7. Let d be a doodle with at least two positive (resp. negative) occurrences of the formula A. The *right contraction* (resp. *left contraction*) *of d on A* is the doodle arising from d by identifying two heads pointing to A (resp. two tails emerging from A). The *maximal contraction* of a doodle d is the result of identifying for every formula A all heads pointing to A and all tails emerging from A.

Example 4.2.8. Let d be the doodle given in (4.3) above. Then the left contraction of d on the formula p is the doodle



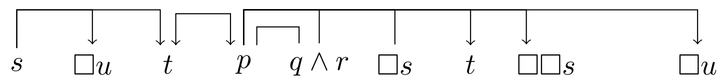
where the superfluous occurrence of the formula p has been omitted. This is also the maximal contraction of d.

Fact 4.2.9. *If a doodle d corresponds to the sequent $\Gamma \Rightarrow \Delta, A, A$, then the right contraction of d on A corresponds to the sequent $\Gamma \Rightarrow \Delta, A$. If the doodle d corresponds to the sequent $B, B, \Sigma \Rightarrow \Pi$, then the left contraction of d on B corresponds to the sequent $B, \Sigma \Rightarrow \Pi$. If the doodle d corresponds to the sequent $\Theta \Rightarrow \Xi$, then the maximal contraction of d corresponds to the sequent* $\mathsf{Supp}\,(\Theta) \Rightarrow \mathsf{Supp}\,(\Xi)$.

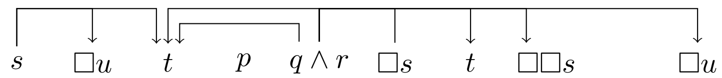
In order to formulate the graphical analogue of cuts between rules we need a notion corresponding to variable elimination for a set of premisses. Of course doodles correspond to sequents instead of premisses, and thus the following definition should be read with variable elimination for a set of sequents in mind. For technical reasons we furthermore maximally contract the resulting doodles.

Definition 4.2.10. Let D be a finite set of doodles and let p be a variable. Then the *p-elimination of D* is the set $D \ominus p$ of doodles containing all maximal contractions of cuts between doodles d_1 and d_2 on p for $d_1, d_2 \in D$ and all doodles from D in which p does not occur.

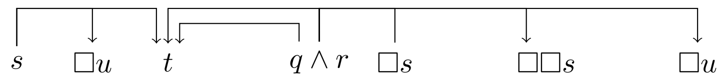
Example 4.2.11. We construct the p-elimination of the four doodles given in (4.4) below:



(4.4)



(4.5)



(4.6)

In a first step we connect the head pointing to the formula p to the tails emerging from p in all possible ways, giving the doodles in (4.5). Then we maximally contract all the doodles, in this case identifying the two heads pointing to the formula t. Together with omitting the now superfluous instances of the formulae p and t this gives the doodles in (4.6).

Remark 4.2.12. Note that a variable p might still occur in the p-elimination of a set of doodles, e.g. if it had two heads of the same doodle pointing at it. In our case this will not be a problem, since this will not occur in the doodles we consider. The definition could be extended to demand multiple cuts completely removing the variable p in this case, but this would necessitate a in this case needless complication to avoid pathological cases as for the set of doodles representing the sequents $\Gamma \Rightarrow \Delta, p, p$ and $p, p, \Sigma \Rightarrow \Pi$.

The connection with the notion of variable elimination for a set of premisses is as follows.

Lemma 4.2.13. *Let $\mathcal{P}$ be a set of premisses and let $\Gamma \Rightarrow \Delta$ be a sequent which does not contain any variables occurring in the premisses $\mathcal{P}$ and such that $\mathsf{Supp}(\Gamma) = \Gamma$ and $\mathsf{Supp}(\Delta) = \Delta$. If D is the set of doodles corresponding to the sequents in $\mathcal{P}(\Gamma \Rightarrow \Delta)$, then for every variable p occurring in $\mathcal{P}$ the set $D \ominus p$ is the set of doodles corresponding to the sequents in $(\mathcal{P} \ominus p)(\Gamma \Rightarrow \Delta)$.*

Proof. We first show that every sequent in $(\mathcal{P} \ominus p)(\Gamma \Rightarrow \Delta)$ is represented by a doodle in $D \ominus p$. Let $(\Sigma \Rightarrow \Pi; \langle F_1, F_2 \rangle)$ be a premiss from $\mathcal{P} \ominus p$ and consider the corresponding sequent

$\Gamma\restriction_{F_1}, \Sigma \Rightarrow \Pi, \Delta\restriction_{F_2}$ from $(\mathcal{P}\ominus p)(\Gamma \Rightarrow \Delta)$. If the sequent is in $\mathcal{P}(\Gamma \Rightarrow \Delta)$, then the variable p does not occur in it. Furthermore, we have a doodle $d \in D$ corresponding to the sequent. But since p does not occur in the doodle d we have $d \in (D \ominus p)$ as well. Suppose on the other hand that there are premisses $(\Sigma_1 \Rightarrow \Pi_1, p; \langle G_1, G_2\rangle)$ and $(p, \Sigma_2 \Rightarrow \Pi_2; \langle H_1, H_2\rangle)$ in $\mathcal{P}$ such that

$$(\Sigma \Rightarrow \Pi; \langle F_1, F_2\rangle) = (\mathsf{Supp}\,(\Sigma_1, \Sigma_2) \Rightarrow \mathsf{Supp}\,(\Pi_1, \Pi_2)\,; \langle G_1 \cup H_1, G_2 \cup H_2\rangle)\ .$$

Then in D we have two doodles d_1 and d_2 corresponding to the two sequents

$$\Gamma\restriction_{G_1}, \Sigma_1 \Rightarrow \Pi_1, p, \Delta\restriction_{G_2} \qquad \text{and} \qquad \Gamma\restriction_{H_1}, p, \Sigma_2 \Rightarrow \Pi_2, \Delta\restriction_{H_2}\ .$$

Thus the cut between the doodles d_1 and d_2 on p corresponds to the sequent

$$\mathsf{Supp}\,(\Gamma\restriction_{G_1}, \Sigma_1, \Gamma\restriction_{H_2}, \Sigma_2) \Rightarrow \mathsf{Supp}\,(\Pi_1, \Delta\restriction_{G_2}, \Pi_2, \Delta\restriction_{H_2})$$

which since $\mathsf{Supp}\,(\Gamma) = \Gamma$ and $\mathsf{Supp}\,(\Delta) = \Delta$ and since the variables in $\Gamma \Rightarrow \Delta$ do not occur in $\Sigma_i \Rightarrow \Pi_i$ for $i = 1, 2$ is the same as the sequent

$$\Gamma\restriction_{G_1 \cup H_1}, \mathsf{Supp}\,(\Sigma_1, \Sigma_2) \Rightarrow \mathsf{Supp}\,(\Pi_1, \Pi_2)\,, \Delta\restriction_{G_2 \cup H_2} \quad = \quad \Gamma\restriction_{F_1}, \Sigma \Rightarrow \Pi, \Delta\restriction_{F_2}\ .$$

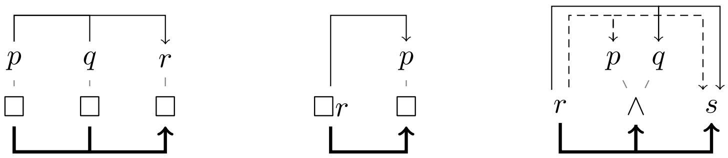
Thus every sequent in $(\mathcal{P}\ominus p)(\Gamma \Rightarrow \Delta)$ is represented by a doodle in $D \ominus p$. Similarly, it is not too difficult to see that every doodle in $D \ominus p$ represents a sequent in $(\mathcal{P}\ominus p)(\Gamma \Rightarrow \Delta)$. $\square$

We would like to extend the graphical representation to rules as well. But since doodles correspond to sequents instead of premisses, we can only represent the premisses instantiated with a particular context. This gives correspondence not directly to rules, but to proto rules instead.

Definition 4.2.14. Let V be a set of variables and let F be a set of formulae such that no variable from V occurs in a formula in F. A *rule doodle* with *context in* F and *principal formulae in* $\Lambda(V)$ is a pair (D, d) where D is a set of doodles over $F \cup V$ and d is a doodle over $F \cup \Lambda(V)$. The doodles in D are the *premisses* and the doodle d is the *conclusion* of the rule doodle. A rule doodle (D, d) *corresponds to* a proto rule $\widehat{R}$ for a rule $R = \mathcal{P}/\Sigma \Rightarrow \Pi$ given by the context $\Gamma \Rightarrow \Delta$ if D is the set of doodles corresponding to the sequents in $\mathcal{P}(\Gamma \Rightarrow \Delta)$ and d is the doodle corresponding to the sequent $\Gamma, \Sigma \Rightarrow \Pi, \Delta$.

In the following we will only consider rule doodles where the conclusion is compatible with the premisses in the sense that in the conclusion there is a head pointing to a formula in the set F of context formulae if and only if in at least one of the premisses there is a head pointing to the same formula, and similarly for tails emerging from formulae in F. To make the construction of the formulae in the conclusion explicit when drawing rule doodles we represent these formulae by their parse trees or in prefix notation. Also, in the first case we often draw the conclusion of a rule doodle at the bottom and its premisses at the top.

Example 4.2.15. The following are three rule doodles with contexts in the sets $\emptyset$, $\{\Box r\}$ and $\{r, s\}$ respectively. Conclusions are drawn thicker at the bottom and premisses at the top.



For the sake of presentation one of the premisses in the last rule doodle is drawn dashed. The rule doodles represent the canonical proto rules for the rules K_2, R_4 and $\wedge_\mathsf{R}$ given by the contexts $\Rightarrow$, $\Box r \Rightarrow$ and $r \Rightarrow s$ respectively.

Even though the context in rule doodles is explicit and thus rule doodles correspond only to proto rules instead of rules, it is clear that every rule with context restrictions can be represented by a rule doodle by encoding the context restrictions in the context of the canonical proto rule and thus in the context of the rule doodle:

Lemma and Definition 4.2.16. *Let $R = \mathcal{P}/\Sigma \Rightarrow \Pi$ be a rule with restrictions and let $\widehat{R}$ be the canonical proto rule for R given by the context $\Gamma \Rightarrow \Delta$ where*

$$\Gamma := \{C \mid C \in F_1 \textit{ for some } \langle F_1, F_2\rangle \textit{ occurring in } \mathcal{P}\}$$
$$\Delta := \{D \mid D \in F_2 \textit{ for some } \langle F_1, F_2\rangle \textit{ occurring in } \mathcal{P}\} \ .$$

Then there is a unique rule doodle corresponding to $\widehat{R}$. We call this the rule doodle for R.

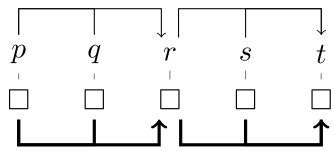
Proof. The doodle (D, d) is constructed by taking D to be the set of (unique) doodles corresponding to the sequents in $\mathcal{P}(\Gamma \Rightarrow \Delta)$ and d to be the (unique) doodle corresponding to the sequent $\Gamma, \Sigma \Rightarrow \Pi, \Delta$. □

The analogue of a cut between rules now takes the following form.

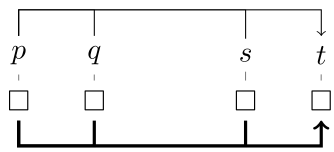
Definition 4.2.17. Let (D_1, d_1) and (D_2, d_2) be two rule doodles with contexts in F_1 resp. F_2 and principal formulae in $\Lambda(V_1)$ resp. $\Lambda(V_2)$ such that no variable occurs in a formula in F_1 and in V_2 or in a formula in F_2 and in V_1 and such that a formula $\heartsuit\vec{p}$ occurs positively in d_1 and negatively in d_2. Then the *cut between* (D_1, d_1) *and* (D_2, d_2) *on* $\heartsuit\vec{p}$ is the rule doodle (D, d) with $D := (D_1 \cup D_2) \ominus \vec{p}$ and where d is the maximal contraction of the cut between the doodles d_1 and d_2 on $\heartsuit\vec{p}$.

While this definition might seem a bit cumbersome, when viewed as a graphical operation a cut between rule doodles is very intuitive: we first identify the two occurrences of the cut formula and then simply pairwise connect all heads pointing to a variable of this formula with all tails emerging from it to construct the new premisses. Analogously connecting heads and tails for the cut formula itself yields the new conclusion.

Example 4.2.18. 1. We first consider a cut between two rule doodles for the rule K_2 in the sense of Lemma 4.2.16. Identifying the two instances of the cut formula $\Box r$ the two rule doodles are

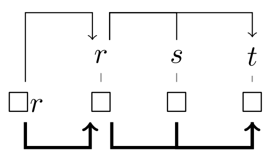


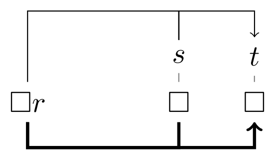
Now computing the cut between the two rules amounts to connecting heads and tails at the node labelled with r and at its parent, giving the rule doodle



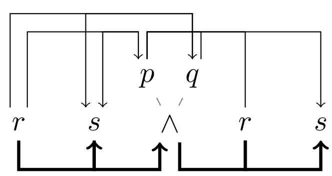
where the superfluous occurrence of $\Box r$ has been omitted.

2. Consider a cut between the rules K_2 and R_4. Again identifying the two occurrences of the cut formula we have the two rule doodles on the left below, and connecting heads and tails at the node labelled with p and its parent node and omitting the superfluous nodes we obtain the rule doodle on the right below, which is easily seen to be the rule doodle for the rule $\mathsf{4}_1$.

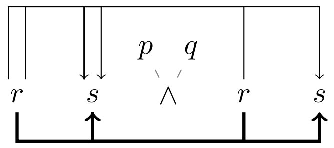


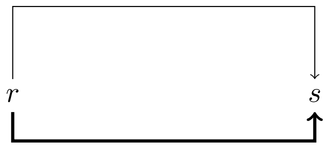


3. A cut between the two rules $\wedge_\mathsf{L}$ and $\wedge_\mathsf{R}$ is constructed in a graphical way by first considering the two rule doodles



Connecting heads and tails at the nodes labelled with p and q and at their parent node gives the diagram on the left below and finally contracting duplicate heads and tails to maximally contract the premisses and omitting the superflous nodes gives the rule doodle on the right below.





The latter is easily identified as the rule doodle for the rule R_{id}.

Indeed, this procedure captures the construction of cuts between (proto) rules:

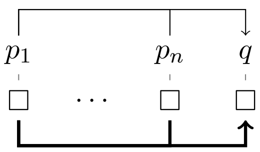
Lemma 4.2.19. *Let $R_1 = \mathcal{P}_1/\Sigma_1 \Rightarrow \Pi_1, \heartsuit\vec{p}$ and $R_2 = \mathcal{P}_2/\heartsuit\vec{p}, \Sigma_2 \Rightarrow \Pi_2$ be two rules and let $\Theta \Rightarrow \Xi$ be a sequent such that the sets $\mathsf{var}\,(\Theta \Rightarrow \Xi)\,, \mathsf{var}\,(\Sigma_1 \Rightarrow \Pi_1)\,, \mathsf{var}\,(\Sigma_2 \Rightarrow \Pi_2)$ are pairwise disjoint. Furthermore, for $i = 1, 2$ let (D_i, d_i) be the rule doodle corresponding to the proto rule for R_i given by the context $\Theta \Rightarrow \Xi$. Then the cut between (D_1, d_1) and (D_2, d_2) on $\heartsuit\vec{p}$ corresponds to the proto rule for the rule $\mathsf{cut}(R_1, R_2, \heartsuit\vec{p})$ given by the context $\mathsf{Supp}\,(\Theta) \Rightarrow \mathsf{Supp}\,(\Xi)$.*

Proof. Since for $i = 1, 2$ the rule doodle (D_i, d_i) corresponds to the proto rule for R_i given by $\Theta \Rightarrow \Xi$, the set $D_1 \cup D_2$ of doodles is the set of doodles corresponding to the sequents in $\mathcal{P}_1(\Theta \Rightarrow \Xi) \cup \mathcal{P}_2(\Theta \Rightarrow \Xi) = (\mathcal{P}_1 \cup \mathcal{P}_2)(\Theta \Rightarrow \Xi)$. Maximally contracting heads pointing to and tails emerging from formulae in Θ, Ξ yields a set D' of doodles corresponding to the sequents in $(\mathcal{P}_1 \cup \mathcal{P}_2)(\mathsf{Supp}\,(\Theta) \Rightarrow \mathsf{Supp}\,(\Xi))$. Thus by repeated applications of Lemma 4.2.13 the set $D := D' \ominus \vec{p}$ is the set of doodles corresponding to the sequents in $((\mathcal{P}_1 \cup \mathcal{P}_2) \ominus \vec{p})(\mathsf{Supp}\,(\Theta) \Rightarrow \mathsf{Supp}\,(\Xi))$. On the other hand, it is clear that the maximal contraction of the cut between the doodles d_1 and d_2 on $\heartsuit\vec{p}$ corresponds to the sequent $\mathsf{Supp}\,(\Theta)\,, \Sigma_1, \Sigma_2 \Rightarrow \mathsf{Supp}\,(\Xi)\,, \Pi_1, \Pi_2$. Thus writing d for this doodle the rule doodle (D, d) corresponds to the proto rule for the rule $\mathsf{cut}(R_1, R_2, \heartsuit\vec{p})$ given by the context $\mathsf{Supp}\,(\Theta) \Rightarrow \mathsf{Supp}\,(\Xi)$. □

Thus in order to construct a cut between two rules we may use the graphical tool to construct the cut between the two rule doodles for the two rules. While the resulting rule doodle corresponds to a proto rule instead of a rule, this proto rule in fact is the canonical proto rule for the cut between the two rules and is easily turned into a rule by turning all formulae occurring both in the premisses and the conclusion into context restrictions in the obvious way.

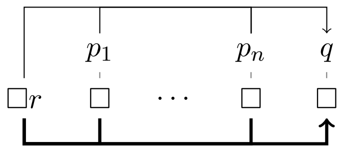
Example 4.2.20. Continuing Example 4.2.18 we have:

1. Repeated cuts between the rule doodles for the rule K_2 yield the rule doodles



which are easily seen to be rule doodles for the rules in the set $\{R_{\mathsf{K}_n} \mid n \geq 2\}$. By construction, this rule set is principal cut closed.

2. Cutting the rule doodle for the rules K_{n+1} with the rule doodle for the rule R_{4} yields the rule doodles



for $n \geq 1$. A second cut with the rule doodle for the rule R_{4} at first adds another tail emerging from the context formula $\Box r$, which is then identified with the original tail. This gives the rule set $\{\mathsf{4}_n \mid n \geq 0\}$ which again by construction is principal cut closed.

3. Since the cut between the two rule doodles for the rules $\wedge_{\mathsf{R}}$ and $\wedge_{\mathsf{L}}$ yields the rule doodle for the identity rule (see Example 4.2.18,3), the rule set $\{\wedge_{\mathsf{R}}, \wedge_{\mathsf{L}}\}$ is principal-cut closed already. Continuing in this fashion for the remaining propositional rules it is not hard to show that the rule sets $\mathsf{G}[\mathsf{cim}]$ are principal-cut closed.

In a similar vein we also have a graphical analogue to the notion of a contraction of a rule.

Definition 4.2.21. Let (D, d) be a rule doodle such that the formulae $\heartsuit\vec{p}$ and $\heartsuit\vec{q}$ both occur positively (resp. negatively) in d. The *right* (resp. *left*) *contraction of* (D, d) *on* $\heartsuit\vec{p}$ *and* $\heartsuit\vec{q}$ is the rule doodle given by (D', d') where D' is constructed from D by renaming every occurrence of a variable in $\vec{q}$ by the corresponding variable in $\vec{p}$ and identifying duplicate heads and tails, and where d' is the doodle constructed from d by renaming $\heartsuit\vec{q}$ to $\heartsuit\vec{p}$ and identifying duplicate heads (resp. tails).

It is clear that this captures the notion of contractions of rules.

Lemma 4.2.22. *Let R be a rule such that the formulae $\heartsuit\vec{p}$ and $\heartsuit\vec{q}$ occur both positively (resp. negatively) in the principal formulae of R. If the rule doodle (D, d) corresponds to the proto rule for R given by the context $\Gamma \Rightarrow \Delta$, then the right (resp. left) contraction of (D, d) on $\heartsuit\vec{p}$ and $\heartsuit\vec{q}$ corresponds to the proto rule for* $\mathsf{ConR}(R, \heartsuit\vec{p}, \heartsuit\vec{q})$ *(resp.* $\mathsf{ConL}(R, \heartsuit\vec{p}, \heartsuit\vec{q})$*) given by the context $\Gamma \Rightarrow \Delta$.*

Proof. Straightforward from the definitions. $\Box$

Thus the contraction of a rule doodle for a rule R is the rule doodle for the contraction of R. This gives us a way of graphically constructing a principal-cut closed *and* contraction closed rule set: we alternate between adding the rules corresponding to cuts between rule doodles and contractions of rule doodles to the rule set. In the easier cases the principal-cut closure of a rule set is almost contraction closed already.

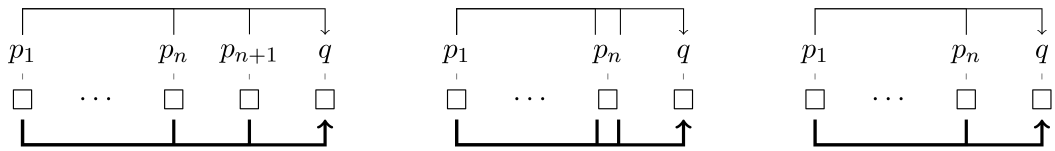
$$\begin{array}{llll} (A1) & \neg\mathbb{C}\top & (A3) & \mathbb{E}A \to A \\ (A2) & \mathbb{E}A \wedge \mathbb{E}B \to \mathbb{E}(A \wedge B) & (A4) & \mathbb{E}A \to \mathbb{C}A \end{array}$$

$$(A5)\ \neg\mathbb{C}\bot$$

$$\mathcal{A}_{\mathsf{ELG1}} := \{(A1), (A2), (A3), (A4)\} \quad \mathcal{A}_{\mathsf{ELG2}} := \mathcal{A}_1 \cup \{(A5)\}$$

Table 4.2: The Hilbert-style axiomatisation for Elgesem's logic of agency and ability.

Example 4.2.23. The rule doodle for the rule K_{n+1} is given below left. In order to construct the left contraction of this rule doodle on $\Box p_n$ and $\Box p_{n+1}$ we first rename the variable p_{n+1} to p_n, which yields the rule doodle in the middle. Finally identifying the tails emerging from p_n resp. $\Box p_n$ gives the rule doodle on the right.



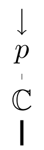
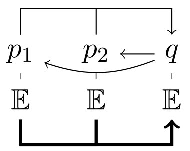
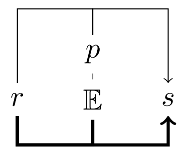
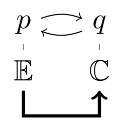
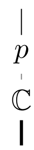
The resulting rule doodle is obviously a rule doodle for the rule K_n. Thus if we want to construct a principal-cut and contraction closed rule set starting only with the rule K_2, then we first add all the cuts as in Example 4.2.20 and then add the missing contraction K_1. It is then easy to see that the resulting rule set is principal-cut closed as well as contraction closed.

The more interesting case is that where adding contractions of rules necessitates closing the rule set under principal-cuts again. The stronger systems of conditional logics considered in the next chapter provide such an involved example. First we apply the graphical method to some less involved examples.

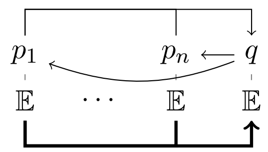
Example 4.2.24 (c)**.** We consider Elgesem's logic of agency and ability [Elg97] and follow [GR05] for the notation and axiomatisation. For the sake of presentation we restrict the framework to only one agent. The extension to multiple agents is straightforward. The logic of agency and ability is based on classical propositional logic with two additional unary modalities $\mathbb{E}$ and $\mathbb{C}$, where $\mathbb{E}A$ is read as "the agent brings it about that A" and $\mathbb{C}A$ is read as "the agent is capable of realising A". The Hilbert-style axiomatisations $\mathcal{A}_{\mathsf{ELG1}}$ and $\mathcal{A}_{\mathsf{ELG2}}$ for the two logics $\mathcal{L}_1$ and $\mathcal{L}_2$ considered in [GR05] are given in Table 4.2. Converting the axioms into rules using the methods of Chapter 3 gives the rules

$$\begin{aligned}
R_{A1} &:= \{(\ \Rightarrow p; \mathcal{C}_\emptyset)\}/\mathbb{C}p \Rightarrow \\
R_{A2} &:= \{(q \Rightarrow p_1; \mathcal{C}_\emptyset), (q \Rightarrow p_2; \mathcal{C}_\emptyset), (p_1, p_2 \Rightarrow q; \mathcal{C}_\emptyset)\}/\mathbb{E}p_1, \mathbb{E}p_2 \Rightarrow \mathbb{E}q \\
R_{A3} &:= \{(p \Rightarrow\ ; \mathcal{C}_{\mathsf{id}})\}/\mathbb{E}p \Rightarrow \\
R_{A4} &:= \{(p \Rightarrow q; \mathcal{C}_\emptyset), (q \Rightarrow p; \mathcal{C}_\emptyset)\}/\mathbb{E}p \Rightarrow \mathbb{C}q \\
R_{A5} &:= \{(p \Rightarrow\ ; \mathcal{C}_\emptyset)\}/\mathbb{C}p \Rightarrow
\end{aligned}$$

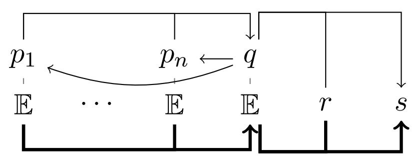
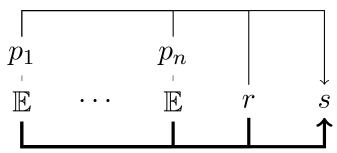
The rule doodles for these rules are shown below.

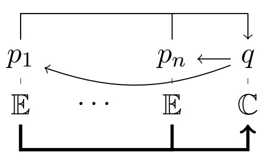
Similar as in the case of the rules K_n repeated cuts on the rule doodle for the rule R_{A2} yield the rule doodles

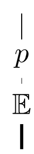


for $n \geq 2$. Call the rules corresponding to these doodles $R_{\mathbb{E}_n}$. Note that these rules are not yet contraction closed, since the left contraction of the rule $R_{\mathbb{E}_2}$, i.e., the rule $R_{\mathbb{E}_1}$, can not be constructed this way. But since contractions of rules preserve soundness of the rule set we may simply add $R_{\mathbb{E}_1}$ to our rule set. Now the cut of the rule doodle for the rule $R_{\mathbb{E}_n}$ with the rule doodle representing the rule R_{A3} as shown on the left below yields the rule doodle shown on the right below, where tautologous premisses, i.e. doodles including circles, are omitted.

Call the rule corresponding to this rule doodle $R_{c\mathbb{E}_n}$. Similarly, a cut between the rule doodle for a rule $R_{\mathbb{E}_n}$ and the one for the rule R_{A4} gives the rule doodle below left, and a cut between the rule doodles for the rules R_{A1} and R_{A4} gives the one below middle.





$$\begin{aligned}
R_{\mathbb{E}_n} &= \{(p_1,\dots,p_n \Rightarrow q; \mathcal{C}_\emptyset)\} \cup \{(q \Rightarrow p_k; \mathcal{C}_\emptyset) \mid 1 \leq k \leq n\} / \mathbb{E}p_1,\dots,\mathbb{E}p_n \Rightarrow \mathbb{E}q \\
R_{c\mathbb{E}_n} &= \{(p_1,\dots,p_n; \mathcal{C}_{\mathsf{id}})\} / \mathbb{E}p_1,\dots,\mathbb{E}p_n \Rightarrow \\
R_{\mathbb{E}\mathbb{C}_n} &= \{(p_1,\dots,p_n \Rightarrow q; \mathcal{C}_\emptyset)\} \cup \{(q \Rightarrow p_k; \mathcal{C}_\emptyset) \mid 1 \leq k \leq n\} / \mathbb{E}p_1,\dots,\mathbb{E}p_n \Rightarrow \mathbb{C}q \\
R_{\mathbb{E}+} &= \{ \Rightarrow p; \mathcal{C}_\emptyset)\} / \mathbb{E}p \Rightarrow \\
R_{A1} &= \{(\Rightarrow p; \mathcal{C}_\emptyset)\} / \mathbb{C}p \Rightarrow \\
R_{A5} &= \{(p \Rightarrow \ \mathcal{C}_\emptyset)\} / \mathbb{C}p \Rightarrow
\end{aligned}$$

$$\begin{aligned}
\mathcal{R}_{\mathsf{ELG1}} &:= \{R_{A1}, R_{\mathbb{E}+}\} \cup \{R_{\mathbb{E}_n} \mid n \geq 1\} \cup \{R_{c\mathbb{E}_n} \mid n \geq 1\} \cup \{R_{\mathbb{E}\mathbb{C}_n} \mid n \geq 1\} \\
\mathcal{R}_{\mathsf{ELG2}} &:= \mathcal{R}_1 \cup \{R_{A5}\}
\end{aligned}$$

Table 4.3: The rules and rule sets for Elgesem's logic of agency and ability

Call the corresponding rules $R_{\mathbb{E}\mathbb{C}_n}$ and $R_{\mathbb{E}+}$ respectively. Analogously for the missing cut between the rules R_{A5} and $R_{\mathbb{E}\mathbb{C}_1}$ we would obtain the additional rule doodle shown above right, but the rule represented by this rule doodle is easily seen to be subsumed by the rule $R_{c\mathbb{E}_1}$. All the constructed rules and the resulting rule sets $\mathcal{R}_{\mathsf{ELG1}}$ and $\mathcal{R}_{\mathsf{ELG2}}$ are given in Table 4.3. An inspection of the corresponding rule doodles shows that both rule sets are principal-cut and contraction closed. Thus since all the rules are shallow by the methods of Chapter 2 we obtain cut elimination and decidability in polynomial space, reproving the semantically driven result for $\mathcal{L}_2$ in [SP08] in a purely syntactic way.

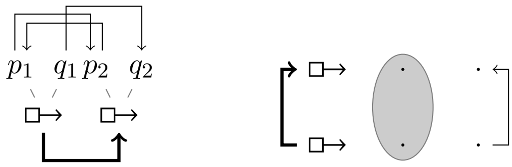
Proposition 4.2.25 (c)**.** *Let $\mathcal{R}_{\mathsf{ELG1}}$ and $\mathcal{R}_{\mathsf{ELG2}}$ be as defined in Table 4.3. Then for $i = 1, 2$ the rule set $\mathcal{R}_{\mathsf{ELG}i}$ is equivalent to the axioms $\mathcal{A}_{\mathsf{ELG}i}$, the sequent system given by $\mathsf{Gc}\mathcal{R}_{\mathsf{ELG}i}\mathsf{Con}$ has cut elimination and its derivability problem is in* PSPACE*. Thus the logics $\mathcal{L}_1$ and $\mathcal{L}_2$ are decidable in polynomial space.*

Proof. Equivalence of the rules and the axioms follows by construction. Since principal-cut closure and contraction closure is easily verified the cut elimination result thus follows from Theorem 2.4.16 and the complexity result follows using the methods of Section 2.7. □

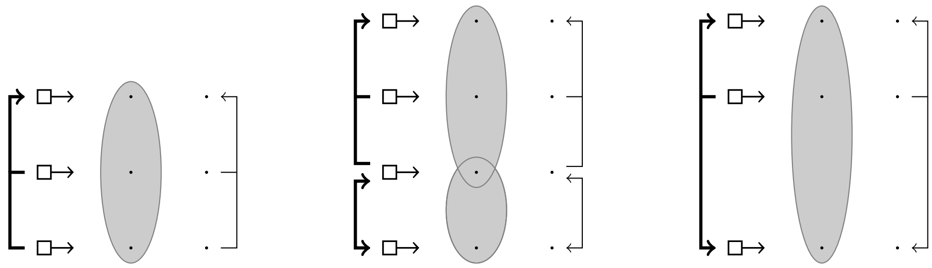
Example 4.2.26 (c)**.** In a way similar to the previous example it is also possible to explicitly construct the rules for the weak conditional logics considered at the end of the previous section in a graphical fashion. In this case it is more convenient to present the formulae in prefix notation instead of as parse trees and to stack them on top of each other. Moreover, for the sake of readability we indicate equivalence of a number of variables by shading all the variables and since the names of the variables are not important we simply replace them with dots. Thus e.g. instead of drawing the rule

$$\{(p_1 \Rightarrow p_2; \mathcal{C}_\emptyset), (p_2 \Rightarrow p_1; \mathcal{C}_\emptyset), (q_1 \Rightarrow q_2; \mathcal{C}_\emptyset)\} / p_1 \mathrel{\Box\!\!\rightarrow} q_1 \Rightarrow p_2 \mathrel{\Box\!\!\rightarrow} q_2$$

corresponding to the axiom (CM) as shown below left we draw it as shown on the right.



In this presentation the axiom (CC) is translated (in the presence of the rule corresponding to (CM)) into the rule doodle shown below left. A cut between this rule doodle and the rule doodle corresponding to the axiom (CEM) is performed as shown below middle and right.



Since all the shaded variables are equivalent, we may simply draw one doodle pointing to the whole block of variables instead of multiple copies of the doodle pointing to each variable. Then again by cutting and contracting rule doodles it is not difficult to construct the rule doodles given in Table 4.4. In particular, note that axiom (ID) translates into the doodle for $R_{\mathsf{CKCEMID}_{0,1}}$, axiom (CS) into the one for $R_{\mathsf{CKCEMCS}_{0,1}}$ and axiom (MP) into that for $R_{\mathsf{CKCEMCS}_{1,0}}$. For any set $\mathcal{A} \subseteq \{(\mathsf{CM}), (\mathsf{CC}), (\mathsf{CN}), (\mathsf{ID}), (\mathsf{CEM}), (\mathsf{MP}), (\mathsf{CS})\}$ with $(\mathsf{CM}), (\mathsf{CC}), (\mathsf{CN}) \in \mathcal{A}$ we then obtain the corresponding rule set as given in Table 4.5, where we denote the set $\mathcal{A}$ by concatenating the axioms and abbreviate CMCCCN to CK. In the graphical representation it is now very easy to see that these rule sets are principal-cut and contraction closed. Then as above we obtain cut elimination and complexity results for all the logics considered. In particular, this gives a graphical reconstruction of the rules sets for weak systems of conditional logic presented in [Che80, PS08, PS09, PS10, PS11].

Proposition 4.2.27 (c). *Let $\mathcal{A} \subseteq \{(\mathsf{CEM}), (\mathsf{CS}), (\mathsf{ID}), (\mathsf{MP})\}$ and let $\mathcal{R}_{\mathsf{CK}\mathcal{A}}$ be the corresponding rule set as defined in Table 4.5. Then $\mathcal{R}_{\mathsf{CK}\mathcal{A}}$ is equivalent to the axioms $\{(\mathsf{CN}), (\mathsf{CM}), (\mathsf{CC})\} \cup \mathcal{A}$, the sequent system given by $\mathsf{Gc}\mathcal{R}_{\mathsf{CK}\mathcal{A}}\mathsf{Con}$ has cut elimination and the derivability problem for $\mathsf{Gc}\mathcal{R}_{\mathsf{CK}\mathcal{A}}$ is in* PSPACE. □

Remark 4.2.28. In the presence of the rules corresponding to the axioms (MP) and (CS) the rules $R_{\mathsf{CKCEMCS}_{n,m}}$ for $n + m > 1$ are actually derivable rules. Thus we might also omit these rules and use the method given in Remark 2.4.18 to show cut elimination. In fact, the rule sets given in [PS09, PS11] have this form.

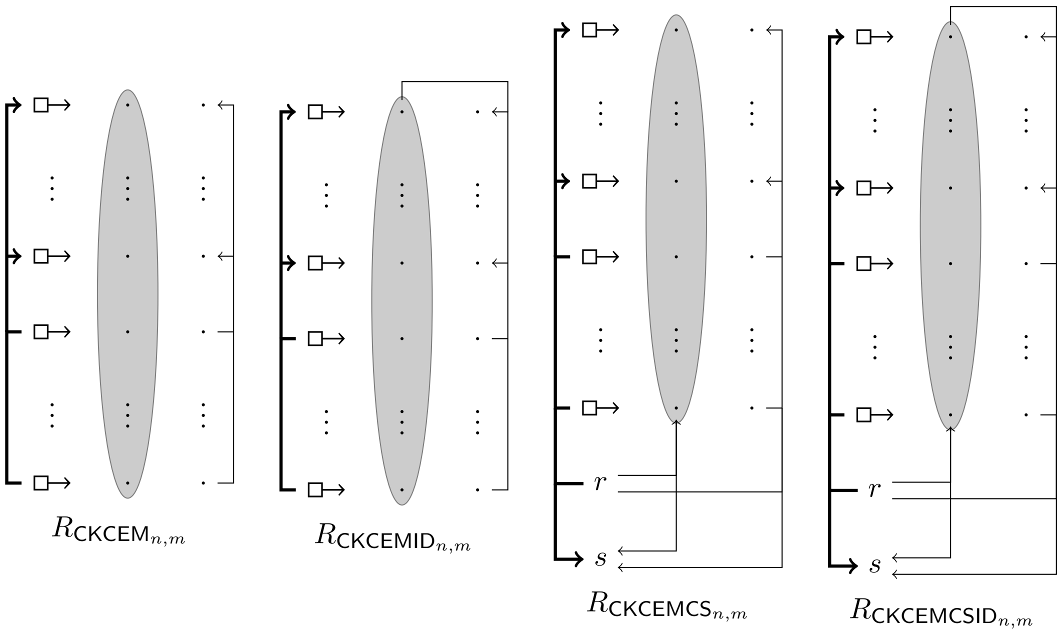


Table 4.4: The rules for the weak conditional logics in their graphical representation. To economise on notation we always take n to be the number of negative and m to be the number of positive literals in the principal formulae.

$$
\begin{aligned}
\mathcal{R}_{\mathsf{CK}} &:= \{R_{\mathsf{CKCEM}_{n,m}} \mid n \geq 0,\ m = 1\}\\
\mathcal{R}_{\mathsf{CKID}} &:= \{R_{\mathsf{CKCEMID}_{n,m}} \mid n \geq 0,\ m = 1\}\\
\mathcal{R}_{\mathsf{CKCEM}} &:= \{R_{\mathsf{CKCEM}_{n,m}} \mid n \geq 0,\ m \geq 1\}\\
\mathcal{R}_{\mathsf{CKIDCEM}} &:= \{R_{\mathsf{CKCEMID}_{n,m}} \mid n \geq 0,\ m \geq 1\}\\
\mathcal{R}_{\mathsf{CKMP}} &:= \mathcal{R}_{\mathsf{CK}} \cup \{R_{\mathsf{CKCEMCS}_{n,m}} \mid n \geq 1,\ m = 0\}\\
\mathcal{R}_{\mathsf{CKIDMP}} &:= \mathcal{R}_{\mathsf{CKID}} \cup \{R_{\mathsf{CKCEMCS}} \mid n \geq 1,\ m = 0\}\\
\mathcal{R}_{\mathsf{CKCEMMP}} &:= \{R_{\mathsf{CKCEMCS}_{n,m}} \mid n + m \geq 1\}\\
\mathcal{R}_{\mathsf{CKIDCEMMP}} &:= \{\mathcal{R}_{\mathsf{CKCEMCSID}_{n,m}} \mid n + m \geq 1\}\\
\mathcal{R}_{\mathsf{CKCS}} &:= \{R_{\mathsf{CKCEMCS}_{n,m}} \mid n \geq 0,\ m = 1\}\\
\mathcal{R}_{\mathsf{CKIDCS}} &:= \{R_{\mathsf{CKCEMCSID}_{n,m}} \mid n \geq 0,\ m = 1\}\\
\mathcal{R}_{\mathsf{CKCEMCS}} &:= \{R_{\mathsf{CKCEMCS}_{n,m}} \mid n \geq 0,\ m \geq 1\}\\
\mathcal{R}_{\mathsf{CKIDCEMCS}} &:= \{R_{\mathsf{CKCEMCSID}_{n,m}} \mid n \geq 0,\ m \geq 1\}\\
\mathcal{R}_{\mathsf{CKMPCS}} &:= \{R_{\mathsf{CKCEMCS}_{n,m}} \mid n + m \geq 1,\ m = 0, 1\}\\
\mathcal{R}_{\mathsf{CKIDMPCS}} &:= \{R_{\mathsf{CKCEMCSID}_{n,m}} \mid n + m \geq 1,\ m = 0, 1\}\\
\mathcal{R}_{\mathsf{CKCEMMPCS}} &:= \{R_{\mathsf{CKCEMCS}_{n,m}} \mid n + m \geq 1\}\\
\mathcal{R}_{\mathsf{CKIDCEMMPCS}} &:= \{R_{\mathsf{CKCEMCSID}_{n,m}} \mid n + m \geq 1\}
\end{aligned}
$$

Table 4.5: The rule sets for weak systems of conditional logic.

Remark 4.2.29. Using these rule sets it is also easy to show that all the logics considered have the Craig interpolation property (see Section 5.5). For the conditional logics axiomatised by $\{(\mathsf{CM}), (\mathsf{CC})\}$ or by $\{(\mathsf{CM}), (\mathsf{CC}), (\mathsf{ID})\}$ this was done in [PS08, PS10]. For most of the other conditional logics this also follows from [PS09, PS11], but was not mentioned there.

4.3 Notes

Principal-cut closure via Cut Trees. The method of constructing a principal-cut closed rule set by absorbing cuts between rules into the rule set most probably was used countless times in the construction of cut-free sequent calculi. But as a systematic procedure it seems not to have been mentioned until [PS08, PS09, PS10, PS11]. Most of our results about small cut trees representing the rules of shallow rule sets for modal logics based on classical propositional logic were originally published in [LP11]. The most pressing open question related to this method is of course whether the resulting rule set can be made contraction closed in a generic and tractable way.

Problem 4.3.1. *Is it possible to use cut trees to automatically construct a tractable, principal-cut closed and contraction closed set of rules from a finite set of shallow rules?*

Combined with the translation from axioms to rules in the classical case this would ensure Pspace-decidability for every logic axiomatised by a finite number of non-iterative axioms and thus give a negative answer to the above mentioned Open Problem 2.8.1 and reduce the complexity bounds from [Lew74, tC05]. Failing this it would also be very interesting to characterise the class of non-iterative axioms for which the principal-cut closure of the translations into rules is contraction closed. Unfortunately, already relatively simple axiom sets seem to produce rules whose principal-cut closure is not contraction closed. One example would be the rule set consisting of the axioms of conditional logic CK together with the axioms (CN') $\bot \mathrel{\Box\!\!\rightarrow} p$, (SDA) $(p \vee q \mathrel{\Box\!\!\rightarrow} r) \to (p \mathrel{\Box\!\!\rightarrow} r)$ and (CA) $(p \mathrel{\Box\!\!\rightarrow} r) \wedge (q \mathrel{\Box\!\!\rightarrow} r) \to (p \vee q \mathrel{\Box\!\!\rightarrow} r)$, which is the basis of the conditional logic S considered in [Che75]. In particular the combination of the axioms (CC) and (CA) seems to be problematic. While this presents a big obstacle to the extension of the automatic method to e.g. stronger conditional logics such as the logic $\mathcal{S}$ of [Bur81] considered e.g. in [SPH10] and under the name PCL in [GGOS09], it also highlights the need for tools to manually construct cut and contraction closed rule sets in case an automatic construction does fail.

Sequent Rules in Pictures The graphical representation of sequent rules developed out of the construction of sequent calculi for the stronger systems of conditional logic considered in the next chapter and is inspired by a talk by Bob Coecke on diagrammatical calculi (see e.g. [Coe10]). The general idea was reported in [LP12a].

There seem to be not too many comparable notions in the literature, the most closely related concept possibly being that of a *logical tomograph* in [Arn08]. This work is based on work by Hertz [Her22, Her23, Her29] and develops the idea of using directed hypergraphs to display the structural relations between atoms in a derivation in classical propositional logic. Sequents with complex formulae are first presented as directed hypergraphs with a root, a directed hyperarc emerging from vertices labelled with the formulae in the antecedent and pointing to the root, and a directed hyperarc emerging from the root and pointing to vertices labelled with the formulae in the succedent. Then successively for each node labelled with a complex formula new nodes labelled with its immediate subformulae are introduced along with new directed hyperarcs depending on the main connective of the formula. Thus e.g. for a vertex v labelled with the formula $A \wedge B$ which is pointed to by a hyperarc two new vertices v_A and v_B are introduced along with two new directed hyperarcs emerging from v and pointing to v_A resp. v_B. Continuing in this fashion the directed hypergraph is turned into the logical tomograph of the sequent, where the external vertices are labelled with propositional variables. Since *maximal strands* in this directed hypergraph then correspond to the sequents in the normal form of the original sequent in the sense of our Definition 3.2.6, this construction can be used to give a graphical decision procedure for classical propositional logic. The main difference of this work to our graphical representation of sequent rules is that the latter is primarily used in the *construction* of a set of sequent rules and not in the actual derivation of a sequent. Of course it is possible to present derivations in a calculus given by rule doodles as sequences of sets of doodles on the underlying set of (parse trees of) formulae, where a set of doodles results from its predecessor in the sequence by an application of a rule doodle. In this way e.g. applications of the propositional rules correspond to certain processes of 'bending and identifying the wires'. But it is not yet clear whether apart from the novelty value this would be of any use.

Concerning other graphical representations considered in proof theory such as proof nets [Gir87] or atomic flows [GG08] the main difference is that the main focus of these methods lies on the investigation of the structure of a whole derivation or on the reduction of unnecessary "bureaucracy", while our method focusses on single rules and serves the much less ambitious goal of providing a tool for the construction of cut-free calculi.

Examples. The logic of agency and ability considered in Example 4.2.24 was originally introduced in [Elg97]. Subsequently in [GR05] it was discovered that the originally proposed axiomatisation is not complete with respect to the given semantics, which led to the introduction of the axiom ($A5$). Our logic $\mathcal{L}_1$ consists of the original axioms, the logic $\mathcal{L}_2$ is the one including the additional axiom. Decidability and PSPACE-complexity results for the logic $\mathcal{L}_2$ were shown in [SP08] by interpreting the semantics in a coalgebraic framework and making use of the there developed generic semantical methods. No sequent calculi seem to have been considered for this logic yet.

Most of the sequent calculi for the weak systems of conditional logic considered in this chapter were known before. For the calculus for the logic CK see e.g. [Che75, Che80]. This calculus and the extension with the axiom (ID) were also considered in [PS08, PS10], where the calculi are moreover used to show that the logics have the Craig interpolation property. Further extensions of the systems with combinations of the axioms $(\mathsf{ID}), (\mathsf{MP}), (\mathsf{CEM})$ were given in [PS09, PS11], where also cut elimination for these systems was shown using a generic syntactic proof. Calculi for extensions of CK including the axiom (CMon) can be found in [SPH10]. There are furthermore a number of results about labelled sequent calculi for these weak systems. A calculus for CK together with a complexity bound has been given in [OS01], and systems for extensions with the axioms $(\mathsf{ID}), (\mathsf{MP}), (\mathsf{CEM}), (\mathsf{CS})$ or the combinations $(\mathsf{ID}) + (\mathsf{MP})$ or $(\mathsf{ID}) + (\mathsf{CEM})$ were introduced in [OPS07], see also [Poz10]. Finally, nested sequent calculi for extensions of CK with combinations of the axioms $\{(\mathsf{ID}), (\mathsf{MP}), (\mathsf{CEM})\}$ or with the axioms (ID) and (CSO) are presented in [AOP12]. These calculi yield PSPACE-decision procedures for the logics in question.

5 Applications: Conditional Logics

In this chapter we will put to use most of the techniques developed in the earlier chapters and construct cut-free sequent calculi and generic decision procedures for a number of stronger systems of conditional logic. The logics we are going to consider are the conditional logics introduced in [Lew73a] which can be characterised in terms of *sphere semantics*. The general interest in these logics stems from the program of giving a formal analysis of the concept of a *counterfactual implication*, i.e. an implication whose antecedent may be false. One of the classical examples for this type of implication (taken from [Lew73a]) is the statement "If kangaroos had no tails, they would topple over". This kind of implication is not really captured by the standard truth-functional material implications, since otherwise due to the fact that the antecedent is false the statement "if kangaroos had no tails, they would not topple over" would need to be true as well, contrary to intuition. On the other hand, a counterfactual implication intuitively should not be antitone in the first argument, since otherwise the statement "if kangaroos had three legs and no tails, they would topple over" would be implied by the first statement. Over the years a number of logics have been proposed to formally capture the notion of counterfactual implication. For an overview see e.g. [NC01, AC09]. In this context we are not going to try to decide which of these systems gives the 'right' analysis of this notion, but we will take this plethora of logics as a suitable testing ground for our generic methods. While these logics can be seen as extensions of the weak conditional logics considered in the previous chapter, we will take the *comparative possibility operator* instead of the (weak or strong) counterfactual implication as the basic operator, since it allows for a simpler and more elegant axiomatisation of conditional logics with sphere semantics. On the other hand since this operator and the (weak or strong) counterfactual implication are interdefinable we retain full expressivity.

After recalling the basic notions and systems of conditional logics with sphere semantics in Section 5.1 we will start by using the translation from Chapter 3 to turn the axioms into rules with context restrictions. From these we will then construct saturated sets of rules using the doodling calculus from the previous chapter, which also give rise to decision procedures for the logics at hand. The main piece of work is done in Section 5.2, where the system for the basic logic $\mathbb{V}_{\preccurlyeq}$ is developed. In Section 5.3 we consider extensions of the logic with additional axioms. By translating the operators the resulting decision procedures also give rise to decision procedures for the logics formulated in the counterfactual implication, which using the circuit presentation of formulae moreover can be seen to be of the same complexity.

Nevertheless, we are also interested in explicit sequent systems for these languages. Thus in Section 5.4 we incorporate the translation between the comparative plausibility operator and the strong counterfactual implication into the rules by adapting the method of cuts between rules and explicitly construct saturated sequent systems for the logics in the language with the latter operator, which again yields decision procedures of matching complexity. Unfortunately, in the case of the weak counterfactual implication the translation is slightly more involved, so as we will see later we cannot use the same method as applied in the case of the strong counterfactual implication. Finally in Section 5.5 we will use the cut-free sequent calculi to show interpolation results for almost all the logics.

5.1 Conditional Logics with Sphere Semantics

We briefly recapitulate the basic notions for conditional logics with sphere semantics from [Lew73a]. The set language of conditional logic is given by the set Λ_{cond} of connectives including apart from the standard boolean connectives the (binary) *comparative plausibility operator* $\preccurlyeq$ and the (binary) *weak* and *strong counterfactual implications* $\Box\!\!\rightarrow$ and $\Box\!\!\Rightarrow$. To save notation we adopt the notational convention that the boolean connectives bind more strongly than the new connectives. Thus we sometimes reduce e.g. the brackets in the formula $((A \vee B) \preccurlyeq (C \wedge D))$ to $(A \vee B \preccurlyeq C \wedge D)$. We read the comparative plausibility operator in a formula $A \preccurlyeq B$ as "formula A is at least as plausible as formula B" and the counterfactual implications in $A \Box\!\!\rightarrow B$ and $A \Box\!\!\Rightarrow B$ as "if A were the case then so would B". The logics considered are based on Lewis' *sphere semantics*:

Definition 5.1.1. Given a set I of worlds, a *system of spheres* over I is a family \$ of subsets of I which is closed under unions and nonempty intersections and which is *nested*, i.e. for every two sets S_1 and S_2 from \$ we have $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$. A *sphere model* then is a triple $\mathcal{I} = (I, (\$_i)_{i \in I}, \pi)$ consisting of a set I of worlds, for every world $i \in I$ an associated system of spheres over I and a valuation $\pi : \mathsf{Var} \to \mathfrak{P}(I)$. The valuation π is extended to a valuation $[\![.]\!] : \mathcal{F}(\Lambda_{\mathsf{cond}}) \to \mathfrak{P}(I)$ on formulae by taking the standard clauses for the boolean connectives in classical propositional logic together with the clauses

$$\begin{aligned}
[\![A \preccurlyeq B]\!] &:= \big\{ i \in I \mid \text{ for all spheres } S \in \$_i \text{ (if } S \cap [\![B]\!] \neq \emptyset \text{ then } S \cap [\![A]\!] \neq \emptyset) \big\} \\
[\![A \Box\!\!\Rightarrow B]\!] &:= \big\{ i \in I \mid \text{ for some sphere } S \in \$_i \, (S \cap [\![A]\!] \neq \emptyset \text{ and } S \subseteq [\![A \to B]\!]) \big\} \\
[\![A \Box\!\!\rightarrow B]\!] &:= [\![A \Box\!\!\Rightarrow B]\!] \cup \big\{ i \in I \mid \bigcup_{S \in \$_i} S \cap [\![A]\!] = \emptyset \big\}
\end{aligned}$$

for the conditional operators.

Intuitively, a system of spheres gives a measure of comparative similarity, worlds in a smaller sphere being more similar to the actual world than those occurring only in larger spheres.

$$
\begin{aligned}
(A \preccurlyeq B) &\leftrightarrow \neg(A \vee B \mathrel{\Box\!\!\Rightarrow} \neg A) \\
&\leftrightarrow \big(A \vee B \mathrel{\Box\!\!\rightarrow} \neg(A \vee B)\big) \vee \neg(A \vee B \mathrel{\Box\!\!\rightarrow} \neg A) \\
(A \mathrel{\Box\!\!\Rightarrow} B) &\leftrightarrow \neg(A \wedge \neg B \preccurlyeq A \wedge B) \\
&\leftrightarrow \neg(A \mathrel{\Box\!\!\rightarrow} \neg A) \wedge (A \mathrel{\Box\!\!\rightarrow} B) \\
(A \mathrel{\Box\!\!\rightarrow} B) &\leftrightarrow (\bot \preccurlyeq A) \vee \neg(A \wedge \neg B \preccurlyeq A \wedge B) \\
&\leftrightarrow \big((A \mathrel{\Box\!\!\Rightarrow} A) \rightarrow (A \mathrel{\Box\!\!\Rightarrow} B)\big)\ .
\end{aligned}
$$

Table 5.1: The equivalences between the conditional connectives

Then a formula $A \preccurlyeq B$ states intuitively that for every B-world there is an A world which is at least as similar to the actual world. Similarly, a formula $A \mathrel{\Box\!\!\Rightarrow} B$ states intuitively that A is considered possible from the point of view of the actual world and that the A-worlds most similar to the actual world are B-worlds as well. The interpretation for a formula $A \mathrel{\Box\!\!\rightarrow} B$ is similar, the only difference being that it is considered true also in case that A is not considered possible from the point of view of the actual world.

A closer analysis of the truth conditions of the different connectives shows that they are interdefinable using the equivalences given in Table 5.1. Thus we may equivalently formulate conditional logics using either of the three connectives. In the following we will take conditional logics $\mathcal{L}$ to be Λ_{cond}-logics based on classical propositional logic in the sense of Definition 2.1.5 and for a given conditional connective $* \in \{\preccurlyeq, \mathrel{\Box\!\!\Rightarrow}, \mathrel{\Box\!\!\rightarrow}\}$ we will write $\mathcal{L}_*$ for the $*$-fragment of $\mathcal{L}$. The logics we are going to consider are defined in terms of classes of sphere models:

Definition 5.1.2. Let $\mathcal{I} = (I, (\$_i)_{i \in I}, \pi)$ be a sphere model. We say that $\mathcal{I}$ is

- *normal* if for every world $i \in I$ we have $\bigcup_{S \in \$_i} S \neq \emptyset$
- *totally reflexive* if for every world $i \in I$ we have $i \in \bigcup_{S \in \$_i} S$
- *weakly centered* if it is normal and for every world $i \in I$ and nonempty sphere $S \in \$_i$ we have $i \in S$
- *centered* if for every world $i \in I$ we have $\{i\} \in \$_i$
- *absolute* if for every two worlds $i, j \in I$ we have $\$_i = \$_j$.

As usual, a formula A is *universally valid* in a sphere model $\mathcal{I} = (I, (\$_i)_{i \in I}, \pi)$ if it holds in every world of $\mathcal{I}$, that is if we have $[\![A]\!] = I$, and it is *universally valid in a class* $\mathcal{C}$ of sphere models if it is universally valid in every sphere model in $\mathcal{C}$. We say that a logic $\mathcal{L}$ is *the logic of a class* $\mathcal{C}$ of sphere models if every formula is a theorem of $\mathcal{L}$ iff it is universally valid in $\mathcal{C}$. Then the logics $\mathbb{V}, \mathbb{VN}, \mathbb{VT}, \mathbb{VW}, \mathbb{VC}, \mathbb{VA}$ and $\mathbb{VNA}$ are defined via

$$
\begin{array}{rl}
\mathsf{CPR} & \dfrac{\vdash B \to A}{\vdash A \preccurlyeq B} \\
(\mathsf{CPA}) & (A \preccurlyeq A \lor B) \lor (B \preccurlyeq A \lor B) \\
(\mathsf{TR}) & (A \preccurlyeq B) \land (B \preccurlyeq C) \to (A \preccurlyeq C) \\
(\mathsf{CO}) & (A \preccurlyeq B) \lor (B \preccurlyeq A)
\end{array}
\qquad
\begin{array}{rl}
(\mathsf{N}) & \neg(\bot \preccurlyeq \top) \\
(\mathsf{T}) & (\bot \preccurlyeq \neg A) \to A \\
(\mathsf{W}) & (\bot \preccurlyeq \neg A) \lor \neg(\neg A \preccurlyeq \top) \to A \\
(\mathsf{C}) & (A \preccurlyeq \top) \land (\top \preccurlyeq A) \to A \\
(\mathsf{A1}) & (A \preccurlyeq B) \to \big(\bot \preccurlyeq \neg(A \preccurlyeq B)\big) \\
(\mathsf{A2}) & \neg(A \preccurlyeq B) \to \big(\bot \preccurlyeq (A \preccurlyeq B)\big)
\end{array}
$$

$$
\mathcal{A}_{\mathbb{V}_{\preccurlyeq}} := \{(\mathsf{CPR}), (\mathsf{CPA}), (\mathsf{TR}), (\mathsf{CO})\}
$$

$$
\begin{array}{lll}
\mathcal{A}_{\mathbb{VN}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{N})\} & \mathcal{A}_{\mathbb{VT}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{T})\} & \mathcal{A}_{\mathbb{VW}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{W})\} \\
\mathcal{A}_{\mathbb{VC}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{C})\} & \mathcal{A}_{\mathbb{VA}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{A1}), (\mathsf{A2})\} & \mathcal{A}_{\mathbb{VNA}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{N}), (\mathsf{A1}), (\mathsf{A2})\}
\end{array}
$$

Table 5.2: Hilbert-style axiomatisations for strong conditional logics formulated using $\preccurlyeq$

- $\mathbb{V}$ is the logic of the class of sphere models
- $\mathbb{VN}$ is the logic of the class of normal sphere models
- $\mathbb{VT}$ is the logic of the class of totally reflexive sphere models
- $\mathbb{VW}$ is the logic of the class of weakly centered sphere models
- $\mathbb{VC}$ is the logic of the class of centered sphere models
- $\mathbb{VA}$ is the logic of the class of absolute sphere models
- $\mathbb{VNA}$ is the logic of the class of normal and absolute sphere models.

If $\mathcal{L}$ is one of these logics, then we also say that a formula A is *$\mathcal{L}$-valid* if A is a theorem of $\mathcal{L}$ and write $\models_{\mathcal{L}} A$.

These logics are known to have sound and complete Hilbert-style axiomatisations, see [Lew73a, Chapter 6]. The equivalent axioms using only the comparative plausibility operator are given in Table 5.2. Using semantical methods it has been established in [FH94] that the validity problem is PSPACE-complete for the logics $\mathbb{V}_{\Box\!\!\to}, \mathbb{VN}_{\Box\!\!\to}, \mathbb{VT}_{\Box\!\!\to}$ and $\mathbb{VC}_{\Box\!\!\to}$ and coNP-complete for the logics $\mathbb{VA}_{\Box\!\!\to}$ and $\mathbb{VNA}_{\Box\!\!\to}$. Since the required resources in these proofs are bounded by the number of subformulae of the input formula, and since the translations given in Table 5.1 increase the number of subformulae only by a constant amount, these complexity results carry over to the other languages as well.

5.2 The Conditional Logic $\mathbb{V}$

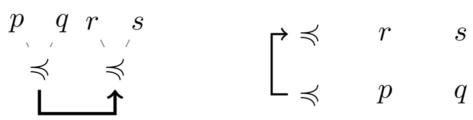
We will now use the methods of Chapters 3 and 4 to construct saturated sets of rules with restrictions first for the logic $\mathbb{V}_{\preccurlyeq}$ and then for its extensions. Note that the Hilbert-style axiomatisation of $\mathbb{V}_{\preccurlyeq}$ as given in Table 5.2 contains the rule CPR as well as axioms. But a straightforward extension of the argument for the equivalence of a Hilbert system given by a set of axioms and the sequent system given by the corresponding set of ground sequents shows that we can translate this rule directly into the sequent rule

$$\mathsf{CPR} \quad \{(q \Rightarrow p; \mathcal{C}_\emptyset)\}/ \ \Rightarrow p \preccurlyeq q$$

(alternatively we could first show that the rule CPR is equivalent to the axiom $A \preccurlyeq (A \wedge B)$ using propositional reasoning and congruence, and then translate this axiom into the rule CPR). Translating the remaining axioms (CPA), (TR) and (CO) into rules using the methods from Section 3.2 yields the rules

$$\begin{array}{rl}
\mathsf{CPA} & \{(q_i \Rightarrow p_1, p_2; \mathcal{C}_\emptyset) \mid i = 1,2\} \cup \{(p_i \Rightarrow q_i; \mathcal{C}_\emptyset) \mid i,j = 1,2\}/ \ \Rightarrow (p_1 \preccurlyeq q_1), (p_2 \preccurlyeq q_2) \\
\mathsf{TR} & \{(p_1 \Rightarrow r; \mathcal{C}_\emptyset), (r \Rightarrow p_1; \mathcal{C}_\emptyset), (q_2 \Rightarrow s; \mathcal{C}_\emptyset), (s \Rightarrow q_2; \mathcal{C}_\emptyset), (q_1 \Rightarrow p_2; \mathcal{C}_\emptyset), (p_2 \Rightarrow q_1; \mathcal{C}_\emptyset)\}/ \\
 & (p_1 \preccurlyeq q_1), (p_2 \preccurlyeq q_2) \Rightarrow (r \preccurlyeq s) \\
\mathsf{CO} & \{(p_1 \Rightarrow q_2; \mathcal{C}_\emptyset), (q_2 \Rightarrow p_2; \mathcal{C}_\emptyset), (p_2 \Rightarrow q_1; \mathcal{C}_\emptyset), (q_1 \Rightarrow p_2; \mathcal{C}_\emptyset)\}/ \ \Rightarrow (p_1 \preccurlyeq q_1), (p_2 \preccurlyeq q_2)
\end{array}$$

Now we make use of the graphical representation of rules as given in the last chapter. For the sake of readability here again we draw the rule doodles in the more economic way seen in Example 4.2.26 by moving the main connective of the formulae to the side and stacking the formulae on top of each other. E.g. instead of drawing the figure on the left below we draw the one on the right.



The rules for conditional logic $\mathbb{V}_{\preccurlyeq}$ then correspond to the rule doodles shown in Figure 5.1. Starting from these rule doodles we now apply the operations of cut (between rule doodles) and contraction (of rule doodles) until we arrive at a saturated rule set, discarding superfluous rule doodles in the process. Note that since all the rules are one-step rules, the rule sets will be automatically context- and mixed-cut closed, and thus we only need to construct a principal-cut and contraction closed rule set. In order to further demonstrate the use of doodles in the construction we consider the steps in detail.

We start by constructing the cut between CPR and TR on the lowermost literal of TR. Omitting the names of variables and marking the variables on which cuts are performed by

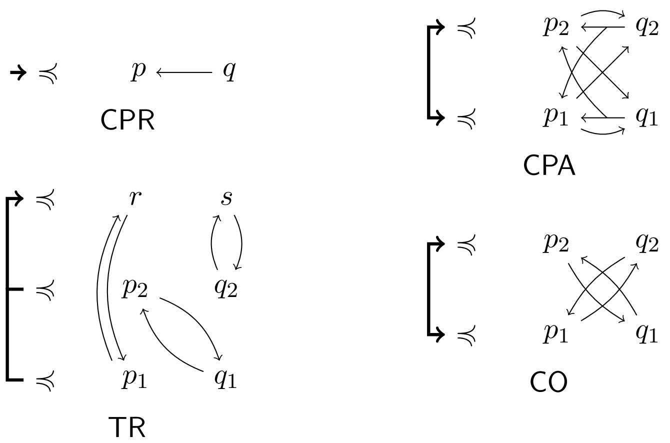
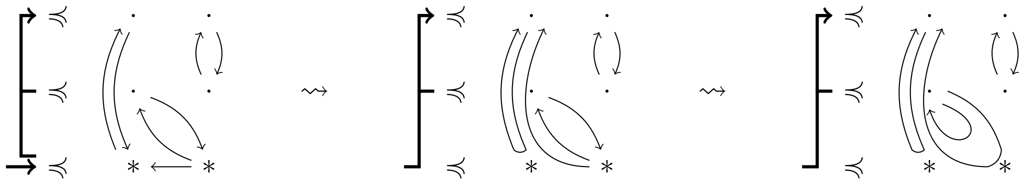
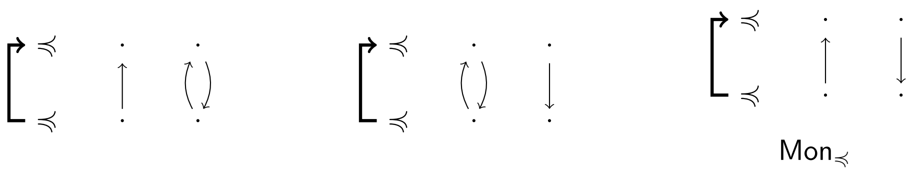


Figure 5.1: The rule doodles corresponding to the translations of the axioms for $\mathbb{V}_{\preccurlyeq}$

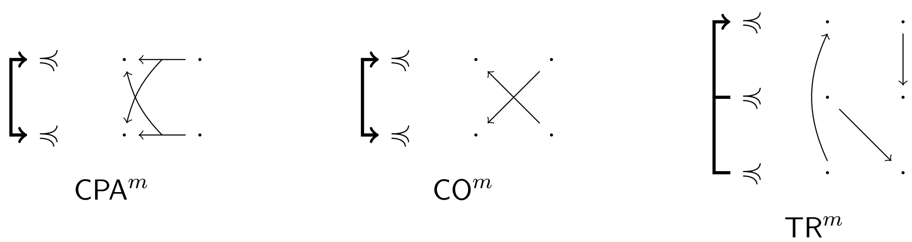
the symbol $*$ we obtain



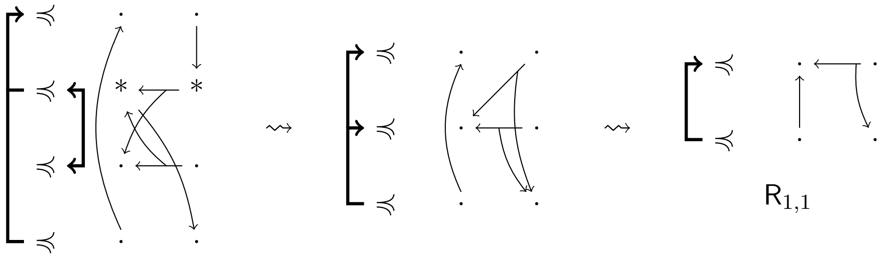
Now yanking the wires, omitting loops (i.e. tautologous premisses) and eliminating superfluous nodes gives the simple rule shown below left. In a similar fashion performing the cut on the other negative literal of TR we construct the rule in the middle below, and performing a cut between these two rules yields the rule $\mathsf{Mon}_{\preccurlyeq}$ shown on the right.



This last rule is particularly useful, since it allows us to greatly simplify the rule set by replacing each rule with its monotone version. This is done by performing cuts between $\mathsf{Mon}_{\preccurlyeq}$ and the rules $\mathsf{CPA}, \mathsf{CO}$ and TR, giving the rules $\mathsf{CPA}^m, \mathsf{CO}^m$ and TR^m shown below.

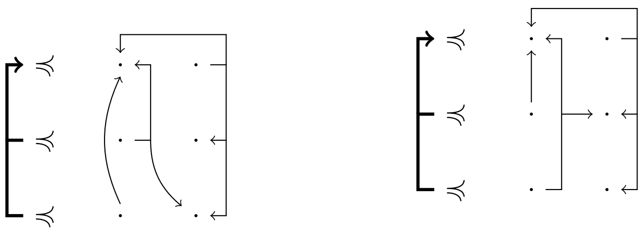


Since each of the original rules is subsumed by their monotone version we may simply omit the original rules from the rule set. Similarly, the new rule CO^m is subsumed by the rule CPA^m, since missing heads of the premisses may be added by weakening. Thus we may omit the former rule as well and continue with the rule set $\{\mathsf{CPR}, \mathsf{Mon}_{\preccurlyeq}, \mathsf{CPA}^m, \mathsf{TR}^m\}$. The next step is to construct a cut between TR^m and CPA^m as shown on the left and middle below (again we mark the variables on which cuts are performed by $*$), followed by a contraction on the two positive literals of the resulting rule, giving the rule shown on the right. Let us write $\mathsf{R}_{1,1}$ for this rule.



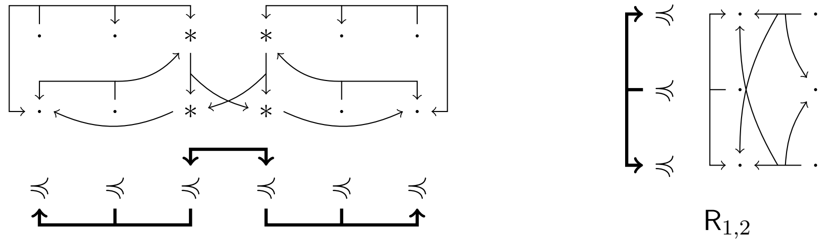
$\mathsf{R}_{1,1}$

Again, since the rule $\mathsf{R}_{1,1}$ subsumes the rule $\mathsf{Mon}_{\preccurlyeq}$ we omit the latter rule from the rule set. Now constructing cuts between TR^m and the new rule $\mathsf{R}_{1,1}$ on the negative literals of TR^m yields the rule on the left below, which we may draw equivalently as shown on the right and which we call $\mathsf{R}_{2,1}$. Note that this rule subsumes the rule TR^m.



$\mathsf{R}_{2,1}$

Performing cuts between two instances of this rule and the rule CPA^m as shown on the left below (rotated) and contracting the two negative literals in the resulting rule gives the rule $\mathsf{R}_{1,2}$ shown on the right below.



$\mathsf{R}_{1,2}$

Now we are almost done. The two rules $\mathsf{R}_{2,1}$ and $\mathsf{R}_{1,2}$ will be the building blocks of our rule set. It is now useful to visualise the rule doodles three dimensionally in a T-shape with the

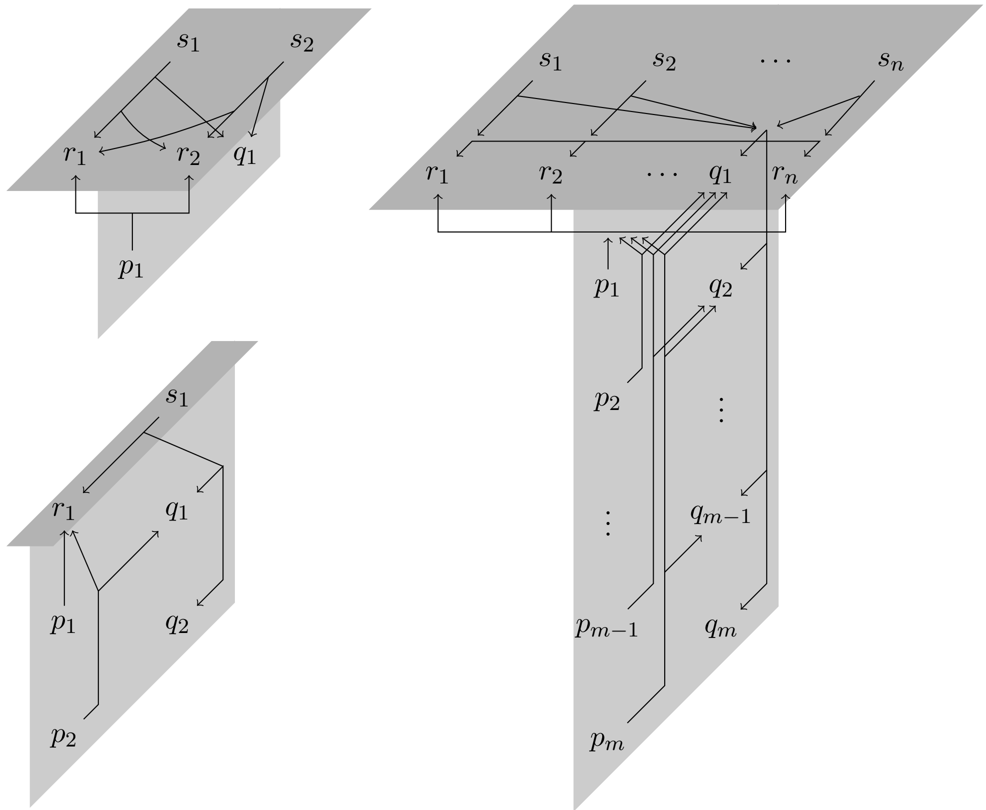
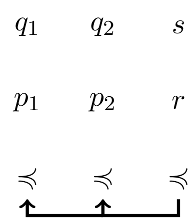
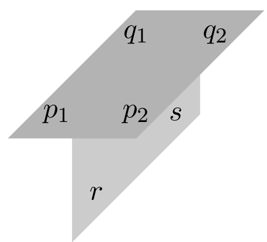


Figure 5.2: The rules for the logic $\mathbb{V}_{\preccurlyeq}$ in graphical representation. The rules $\mathsf{R}_{1,2}$ and $\mathsf{R}_{2,1}$ are shown on the left and the general rule $\mathsf{R}_{m,n}$ is shown on the right.

positive literals in the top bar and the negative literals in the stem of the T. Furthermore, since this implicitly encodes the principal formulae, we omit the nodes labelled with $\preccurlyeq$ and the corresponding doodle. Thus e.g. instead of drawing the figure on the left below we draw the one on the right.





Then the rules $\mathsf{R}_{1,2}$ and $\mathsf{R}_{2,1}$ are drawn as shown on the left in Figure 5.2. Now repeatedly cutting instances of the rule $\mathsf{R}_{1,2}$ yields for $n \geq 2$ a rule $\mathsf{R}_{1,n}$ with n positive literals and one negative literal in the principal formulae. Similarly, cutting instances of $\mathsf{R}_{2,1}$ on the lowermost negative literals yields for $m \geq 2$ a rule $\mathsf{R}_{m,1}$ with m negative literals and one positive literal in the principal formulae. Now cutting these two rules on the negative literal of the former and the positive literal of the latter gives the rule $\mathsf{R}_{m,n}$ with m negative and n positive literals in the principal formulae as shown in Figure 5.2 on the right. Here in order to enhance readability we draw some doodles with a head pointing to another doodle instead of drawing heads pointing to all the formulae occurring positively in the latter. The following definition captures these rules in non-graphical notation.

Definition 5.2.1. For $m \geq 0, n \geq 1$ the rule $\mathsf{R}_{m,n}$ is defined as the rule $\mathcal{P}_{m,n}/\Sigma_m \Rightarrow \Pi_n$ where

$$\begin{aligned} \mathcal{P}_{m,n} &:= \{(s_k \Rightarrow q_1, \dots, q_m, r_1, \dots, r_n; \mathcal{C}_\emptyset) \mid 1 \leq k \leq n\} \\ &\quad \cup \{(p_k \Rightarrow q_1, \dots, q_{k-1}, r_1, \dots, r_n; \mathcal{C}_\emptyset) \mid 1 \leq k \leq m\} \\ \Sigma_m \Rightarrow \Pi_n &:= (p_1 \preccurlyeq q_1), \dots, (p_m \preccurlyeq q_m) \Rightarrow (r_1 \preccurlyeq s_1), \dots, (r_n \preccurlyeq s_n)\ . \end{aligned}$$

Furthermore, define $\mathcal{R}_{\mathbb{V}_\preccurlyeq} := \{\mathsf{R}_{m,n} \mid m \geq 0, n \geq 1\}$.

The rules in $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$ are sound for the logic $\mathbb{V}_\preccurlyeq$ by construction, and since they subsume each of the original rules we get completeness (with Cut) as well. This can also be seen directly.

Theorem 5.2.2 (c)**.** *The sequent calculus $\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{CutCon}$ is sound and complete for conditional logic $\mathbb{V}_\preccurlyeq$.*

Proof. For soundness we show that if $\Sigma \Rightarrow \Pi$ is derivable in $\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{CutCon}$, then the formula $\bigwedge \Sigma \to \bigvee \Pi$ is valid in all models of $\mathbb{V}_\preccurlyeq$. As usual this follows from showing that whenever for every premiss $\Gamma \Rightarrow \Delta$ of an application of a rule in $\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{CutCon}$ the formula $\bigwedge \Gamma \to \bigvee \Delta$ is valid in a model of $\mathbb{V}_\preccurlyeq$, then so is the conclusion of this application. The proof for the propositional rules as well as for Cut and Con is standard.

So suppose that for some $m \geq 0,\ n \geq 1$ the last applied rule was $R_{m,n}$, with conclusion $(A_1 \preccurlyeq B_1), \dots, (A_m \preccurlyeq B_m) \Rightarrow (C_1 \preccurlyeq D_1), \dots, (C_n \preccurlyeq D_n)$ and premisses

$$\{D_k \Rightarrow B_1, \dots, B_m, C_1, \dots, C_n \mid 1 \leq k \leq n\} \cup \{A_k \Rightarrow B_1, \dots, B_{k-1}, C_1, \dots, C_n \mid 1 \leq k \leq m\}$$

Furthermore, suppose all the premisses are $\mathbb{V}_\preccurlyeq$-valid. Let $\mathcal{I} = (I, (\$_i)_{i \in I}, \pi)$ be a sphere model and take an arbitrary world $i \in I$. We need to show that the formula $\bigwedge_{k=1}^m (A_k \preccurlyeq B_k) \to \bigvee_{\ell=1}^m (C_\ell \preccurlyeq D_\ell)$ holds at i. So suppose $i \in [\![A_k \preccurlyeq B_k]\!]$ for all $k \leq m$ and that for a $k \leq n$ we have $i \notin [\![C_\ell \preccurlyeq D_\ell]\!]$ for all ℓ with $1 \leq \ell \leq n$ and $\ell \neq k$. If $\bigcup \$_i \cap [\![D_k]\!] = \emptyset$, then we have $i \in [\![C_k \preccurlyeq D_k]\!]$ and are done. In particular, this is the case if $\bigcup \$_i = \emptyset$. Otherwise, choose a sphere $S \in \$_i$ and a world $j \in S \cap [\![D_k]\!]$. Since the premisses of the rule application are $\mathbb{V}_\preccurlyeq$-valid, we have $\models_{\mathbb{V}_\preccurlyeq} D_k \to \bigvee_{\ell=1}^m B_\ell \vee \bigvee_{\ell=1}^n C_\ell$, which gives $j \in \bigcup_{\ell=1}^m [\![B_\ell]\!] \cup \bigcup_{\ell=1}^n [\![C_\ell]\!]$. Thus $j \in \bigcup_{\ell=1}^n [\![C_\ell]\!]$ or $j \in [\![B_\ell]\!]$ for some $\ell \leq m$. In the latter case, since $i \in [\![A_\ell \preccurlyeq B_\ell]\!]$ we find a $j_2 \in S \cap [\![A_\ell]\!]$, and since $\models_{\mathbb{V}_\preccurlyeq} A_\ell \to \bigvee_{\ell'=1}^{\ell-1} B_{\ell'} \vee \bigvee_{\ell'=1}^n C_{\ell'}$ we find a world $j_2 \in \bigcup_{\ell'=1}^{\ell-1} [\![B_{\ell'}]\!] \cup \bigcup_{\ell'=1}^n [\![C_{\ell'}]\!]$. Continuing like this we find a world $j' \in I$

with $j' \in S \cap \bigcup_{\ell=1}^{n} [\![C_\ell]\!]$. Now if $j' \notin [\![C_k]\!]$, then there is a $\ell \neq k$ with $j' \in [\![C_\ell]\!]$. But since $i \notin [\![C_\ell \preccurlyeq D_\ell]\!]$ there is a sphere $S' \in \$_i$ with $S' \subsetneqq S$ and $S' \cap [\![D_\ell]\!] \neq \emptyset$ and $S' \cap [\![C_\ell]\!] = \emptyset$. As above we find a world $j'' \in S' \cap \bigcup_{t=1}^{n} [\![C_t]\!] = (S' \cap \bigcup_{t=1}^{n} [\![C_t]\!]) \setminus [\![C_\ell]\!]$. Repeating the argument we finally obtain a sphere $S'' \in \$_i$ with $\emptyset \neq S'' \cap \bigcup_{\ell=1}^{n} [\![C_\ell]\!] = S'' \cap [\![C_k]\!]$, and since by construction $S'' \subseteq S$ we have $i \in [\![C_k \preccurlyeq D_k]\!]$. Thus the formula $\bigwedge_{k=1}^{m}(A_k \preccurlyeq B_k) \to \bigvee_{\ell=1}^{m}(C_\ell \preccurlyeq D_\ell)$ holds at world i.

For completeness by the completeness result for $\mathcal{H}\mathsf{c}\mathcal{A}_{\mathbb{V}_\preccurlyeq}$ in [Lew73a, Chapter 6] it suffices to show that all axioms and rules of $\mathcal{H}\mathsf{c}\mathcal{A}_{\mathbb{V}_\preccurlyeq}$ are derivable in $\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{CutCon}$. But this is easy: rule CPR is derived using $\mathsf{R}_{0,1}$, axioms (CPA) and (CO) using rule $\mathsf{R}_{0,2}$ and axiom (TR) using rule $\mathsf{R}_{2,1}$. □

Furthermore, almost by construction we obtain saturation of the rule set.

Theorem 5.2.3 (c)**.** *The rule set* $\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}$ *is saturated.*

Proof. The rules in $\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}$ are all shallow rules and thus automatically mixed- and context-cut closed. Furthermore, since the rule set Gc is saturated (see Example 2.4.14) and since the rules in $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$ mention only the conditional connectives, we only need to check principal-cut closure and contraction closure for the rules in $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$. For contraction closure it is fairly obvious from the graphical representation of the rules that the right contraction of a rule $\mathsf{R}_{m,n+1}$ and the left contraction of a rule $\mathsf{R}_{m+1,n}$ are subsumed by the rule $\mathsf{R}_{m,n}$. The main idea for showing that the rule set $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$ is principal-cut closed is to show that cuts between the rules $\mathsf{R}_{2,1}$ and $\mathsf{R}_{1,2}$ of the form as shown in Figure 5.3 top left and middle are subsumed by a cut of the form shown on the top right. But this can be seen easily by weakening the appropriate premiss. Since a rule $\mathsf{R}_{m,n}$ can be viewed as the result of performing cuts between $m-1$ instances of $\mathsf{R}_{2,1}$ and $n-1$ instances of $\mathsf{R}_{1,2}$, we can use this to permute all the instances of $\mathsf{R}_{1,2}$ in a cut between two rules $\mathsf{R}_{m,n}$ and $\mathsf{R}_{k,\ell}$ to the top. Moreover, cuts between two instances of the rule $\mathsf{R}_{2,1}$ of the form shown in Figure 5.3 bottom left are similarly easily seen to be subsumed by cuts of the form shown on the right. We use this to successively permute instances of $\mathsf{R}_{2,1}$ so that cuts are only performed on the lowest negative literal. Intuitively this amounts to 'straightening the stem of the T'. Since multiple cuts between the rules $\mathsf{R}_{2,1}$ and $\mathsf{R}_{1,2}$ of this form are exactly the rules in $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$, this shows that a cut between two rules $\mathsf{R}_{m,n}$ and $\mathsf{R}_{k,\ell}$ is subsumed by the rule $\mathsf{R}_{m+k-1,n+\ell-1}$. □

It is worth noting that the proof of Theorem 5.2.3 also shows that the cut closure of a small set of rules generating all the rules in $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$ is contraction closed.

Corollary 5.2.4 (c)**.** *Let* $\mathcal{R}'_{\mathbb{V}_\preccurlyeq} := \{\mathsf{R}_{0,1}, \mathsf{R}_{1,1}, \mathsf{R}_{2,1}, \mathsf{R}_{1,2}\}$. *Then* $\mathsf{Gc}\mathcal{R}'_{\mathbb{V}_\preccurlyeq}\mathsf{CutCon}$ *is sound and complete for* $\mathbb{V}_\preccurlyeq$ *and the rule set* $\mathsf{cc}\left(\mathcal{R}'_{\mathbb{V}_\preccurlyeq}\right)$ *is contraction closed.*

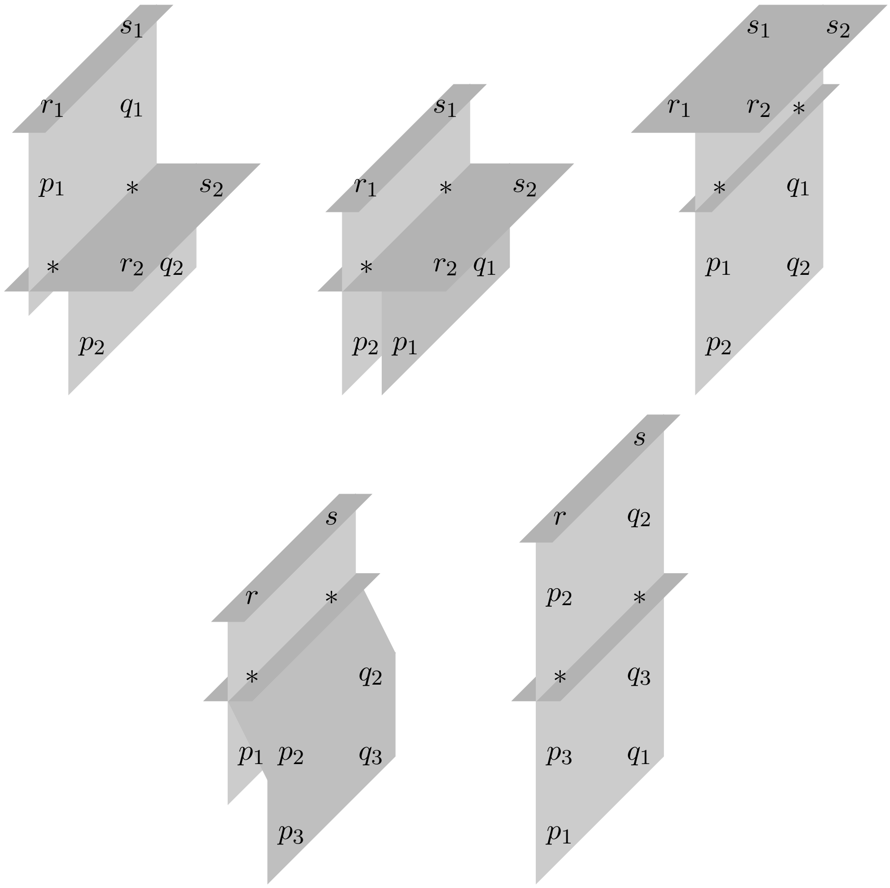


Figure 5.3: Straightening the rules in the proof of Theorem 5.2.3

Proof. Soundness and completeness for $\mathbb{V}_{\preccurlyeq}$ follow from the proof of Theorem 5.2.2. Contraction closure of the cut closure follows by bringing cuts between rules in the form corresponding to rules in $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$ as in the proof of Theorem 5.2.3 and using contraction closure of $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$. □

Thus making use of either the generic cut elimination theorem and decision procedures from Chapter 2 for $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$ or the results about small representations of rules in the cut closure of $\mathcal{R}'_{\mathbb{V}_{\preccurlyeq}}$ we have cut elimination and a complexity result for $\mathbb{V}_{\preccurlyeq}$.

Corollary 5.2.5 (c)**.** *The rule set* $\mathsf{Gc}\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathsf{Con}$ *has cut elimination and thus the validity problem for* $\mathbb{V}_{\preccurlyeq}$ *is in* PSPACE*. More precisely, it is solvable in space polynomial in the circuit size of the input.*

Proof. Since the rule set is saturated, Theorem 2.4.16 yields cut elimination. Furthermore, it is not hard to see that the rule set $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$ is tractable. Thus using Corollary 2.7.9 and the fact that $\mathsf{Gc}\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathsf{CutCon}$ is sound and complete for $\mathbb{V}_{\preccurlyeq}$ we obtain the complexity result. Equivalently, the latter result follows from Corollaries 4.1.21 and 5.2.4 using the representation of rules in the cut closure of $\mathcal{R}'_{\mathbb{V}_{\preccurlyeq}}$ as small cut trees. □

Remark 5.2.6. In the spirit of [Sch07] we may now also use the translation from rules to axioms to give an alternative axiomatisation for the logic $\mathbb{V}_{\preccurlyeq}$. Since the two rules $\mathsf{R}_{0,1}$ and $\mathsf{R}_{2,2}$ generate all the rules in $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$ using cuts and contractions, it suffices to translate these two rules into axioms. Instead of using the translation from Section 3.3 in this case due to the relatively simple structure of the premisses it is easier to translate the rules directly. We do this by identifying a variable q which occurs in only one premiss and substitute the variable according to this premiss. For the rule $\mathsf{R}_{0,1}$ we have the premiss $q \Rightarrow p$ and the conclusion $p \preccurlyeq q$, and it is clear that by substituting $p \wedge r$ for q we force the implication $p \wedge r \rightarrow p$ of the premiss and obtain the equivalent axiom $(p \wedge r) \mathrel{\Box\!\!\rightarrow} p$. For the rule $\mathsf{R}_{2,2}$ we apply the same method and obtain the equivalent axiom

$$((r_1 \vee r_2) \wedge t_1 \preccurlyeq q_1) \wedge ((r_1 \vee r_2) \wedge t_2 \preccurlyeq q_2) \rightarrow (r_1 \preccurlyeq (r_1 \vee r_2 \vee q_1 \vee q_2) \wedge u_1) \vee (r_2 \preccurlyeq (r_1 \vee r_2 \vee q_1 \vee q_2) \wedge u_2)\ .$$

While the Hilbert-system given by these two axioms then is sound and complete for $\mathbb{V}_{\preccurlyeq}$, the original axiomatisation of course is slightly superior in terms of readability and intuitiveness.

5.3 Extensions of $\mathbb{V}$

We now extend the results of the previous section to the extensions of the logic $\mathbb{V}$ mentioned in Section 5.1. The logics $\mathbb{VN}, \mathbb{VT}, \mathbb{VW}, \mathbb{VC}, \mathbb{VA}$ and $\mathbb{VNA}$ arise by extending the axiomatisation of $\mathbb{V}$ with additional axioms from

(N)	$\neg(\bot \preccurlyeq \top)$	(C)	$((A \preccurlyeq \top) \wedge (\top \preccurlyeq A)) \rightarrow A$
(T)	$(\bot \preccurlyeq \neg A) \rightarrow A$	(A1)	$(A \preccurlyeq B) \rightarrow (\bot \preccurlyeq \neg(A \preccurlyeq B))$
(W)	$((\bot \preccurlyeq \neg A) \vee \neg(\neg A \preccurlyeq \top)) \rightarrow A$	(A2)	$\neg(A \preccurlyeq B) \rightarrow (\bot \preccurlyeq (A \preccurlyeq B))$

in the way as given in Table 5.2, which we repeat here for convenience:

$\mathcal{A}_{\mathbb{VN}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{N})\}$	$\mathcal{A}_{\mathbb{VT}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{T})\}$	$\mathcal{A}_{\mathbb{VW}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{W})\}$
$\mathcal{A}_{\mathbb{VC}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{C})\}$	$\mathcal{A}_{\mathbb{VA}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{A1}), (\mathsf{A2})\}$	$\mathcal{A}_{\mathbb{VNA}_{\preccurlyeq}} := \mathcal{A}_{\mathbb{V}_{\preccurlyeq}} \cup \{(\mathsf{N}), (\mathsf{A1}), (\mathsf{A2})\}$

Note that apart from the axioms (A1) and (A2) all the axioms are non-iterative. The axioms (N) and (T) are already translatable clauses in the sense of Definition 3.2.9 extended as by Remark 3.2.11. Thus we can translate them immediately using the methods of Chapter 3. This yields the two rules

$$\mathsf{R}_{\mathsf{N}} := \{(p \Rightarrow\ ; \mathcal{C}_{\emptyset}), (\ \Rightarrow q; \mathcal{C}_{\emptyset})\}/(p \preccurlyeq q) \Rightarrow$$
$$\mathsf{R}_{\mathsf{T}} := \{(p \Rightarrow\ ; \mathcal{C}_{\emptyset}), (\ \Rightarrow q; \mathcal{C}_{\mathsf{id}}\}/(p \preccurlyeq q) \Rightarrow \quad .$$

$$\begin{aligned}
\mathsf{R_N} &:= \{(p \Rightarrow\ ; \mathcal{C}_\emptyset), (\ \Rightarrow q; \mathcal{C}_\emptyset)\}/(p \preccurlyeq q) \Rightarrow \\
\mathsf{R_T} &:= \{(p \Rightarrow\ ; \mathcal{C}_\emptyset), (\ \Rightarrow q; \mathcal{C}_{\mathsf{id}}\}/(p \preccurlyeq q) \Rightarrow \quad . \\
\mathsf{R_W} &:= \{(\ \Rightarrow p; \mathcal{C}_{\mathsf{id}}), (\ \Rightarrow q; \mathcal{C}_\emptyset)\}/\ \Rightarrow (p \preccurlyeq q) \\
\mathsf{R_C} &:= \{(p \Rightarrow\ ; \mathcal{C}_{\mathsf{id}}), (\ \Rightarrow q; \mathcal{C}_\emptyset)\}/(p \preccurlyeq q) \Rightarrow \quad . \\
\mathsf{R_A} &:= \{(p \Rightarrow\ ; \mathcal{C}_\emptyset), (q \Rightarrow\ ; \mathcal{C}_{\mathbb{VA}})\}/\ \Rightarrow (p \preccurlyeq q)
\end{aligned}$$

Table 5.3: The translations of the additional axioms for extensions of $\mathbb{V}_\preccurlyeq$

The axioms (W) and (C) can be simplified first. By propositional reasoning adding axiom (W) is equivalent to adding axiom (T) and the translatable clause

$$(\mathsf{W}') \quad \neg(\neg A \preccurlyeq \top) \to A \ .$$

Furthermore, since using rule CPR the formula $(\top \preccurlyeq A)$ is derivable in $\mathbb{V}$ for every formula A, we may replace axiom (C) with the translatable clause

$$(\mathsf{C}') \quad (A \preccurlyeq \top) \to A \ .$$

Translating these two axioms into rules then yields the rules

$$\begin{aligned}
\mathsf{R_W} &:= \{(\ \Rightarrow p; \mathcal{C}_{\mathsf{id}}), (\ \Rightarrow q; \mathcal{C}_\emptyset)\}/\ \Rightarrow (p \preccurlyeq q) \\
\mathsf{R_C} &:= \{(p \Rightarrow\ ; \mathcal{C}_{\mathsf{id}}), (\ \Rightarrow q; \mathcal{C}_\emptyset)\}/(p \preccurlyeq q) \Rightarrow \quad .
\end{aligned}$$

For the axioms ($\mathsf{A1}$) and ($\mathsf{A2}$) the translation is not quite as straightforward. But using methods as in the case of the axioms for $\mathsf{S5}$ we obtain the rule

$$\mathsf{R_A} := \{(p \Rightarrow\ ; \mathcal{C}_\emptyset), (q \Rightarrow\ ; \mathcal{C}_{\mathbb{VA}})\}/\ \Rightarrow (p \preccurlyeq q)$$

where $\mathcal{C}_{\mathbb{VA}}$ is the context restriction $\langle\{r \preccurlyeq s\}, \{r \preccurlyeq s\}\rangle$. The details of the construction are given below. These newly constructed rules are summarised in Table 5.3. Then by construction we obtain soundness and completeness of the sequent calculi with the cut rule.

Proposition 5.3.1 (c)**.** *The following calculi are sound and complete for the specified logics:*

$\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R_N}\mathsf{CutCon}$	*for* $\mathbb{VN}_\preccurlyeq$	$\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R_T}\mathsf{CutCon}$	*for* $\mathbb{VT}_\preccurlyeq$
$\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R_T}\mathsf{R_W}\mathsf{CutCon}$	*for* $\mathbb{VW}_\preccurlyeq$	$\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R_C}\mathsf{CutCon}$	*for* $\mathbb{VC}_\preccurlyeq$
$\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R_A}\mathsf{CutCon}$	*for* $\mathbb{VA}_\preccurlyeq$	$\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R_N}\mathsf{R_A}\mathsf{CutCon}$	*for* $\mathbb{VNA}_\preccurlyeq$

Proof. From the completeness results in [Lew73a, Chapter 6], by first translating the axioms given there into the language using only the comparative plausibility operator using the equivalences given in Table 5.1 we obtain soundness and completeness of the Hilbert-systems as specified in Table 5.2. Then reasoning as above and using Corollary 3.2.26 we can show that the constructed rules are equivalent to the axioms over $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathsf{Mon}_{\preccurlyeq}$, where $\mathsf{Mon}_{\preccurlyeq}$ $\{(p \Rightarrow r; \mathcal{C}_{\emptyset}), (s \Rightarrow q; \mathcal{C}_{\emptyset})\}/(p \preccurlyeq q) \Rightarrow (r \preccurlyeq s)$ is the rule constructed on page 147. Since this rule is subsumed by the rule $\mathsf{R}_{1,1}$ from $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$, the constructed rules evidently are equivalent to the axioms also over $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$. This gives the result for the systems not including the rule R_{A}.

For the systems including the rule R_{A} we first show that adding both axioms ($\mathsf{A1}$) and ($\mathsf{A2}$) is equivalent over $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$ to adding the ω-set

$$\left\{ \bigwedge_{i=1}^{n}(r_i \preccurlyeq s_i) \wedge \neg \bigvee_{j=1}^{m}(t_j \preccurlyeq u_j) \to \left(\bot \preccurlyeq \neg \bigwedge_{i=1}^{n}(r_i \preccurlyeq s_i) \vee \bigvee_{j=1}^{m}(t_j \preccurlyeq u_j) \right) \mid n, m \geq 0 \right\}$$

for the axiom

$$(\mathsf{A}) \quad (r \preccurlyeq s) \wedge \neg(t \preccurlyeq u) \to (\bot \preccurlyeq \neg(r \preccurlyeq s) \vee (t \preccurlyeq u)) \ .$$

For the one direction it is clear that both ($\mathsf{A1}$) and ($\mathsf{A2}$) are logically equivalent (modulo injective renaming) to formulae in the ω-set. For the other direction, if we have axioms ($\mathsf{A1}$) and ($\mathsf{A2}$), then we can derive for every i, j with $1 \leq i \leq n$ and $1 \leq j \leq m$ the formulae

$$(r_i \preccurlyeq s_i) \to (\bot \preccurlyeq \neg(r_i \preccurlyeq s_i)) \quad \text{and} \quad \neg(t_j \preccurlyeq u_j) \to (\bot \preccurlyeq (t_j \preccurlyeq u_j))$$

and thus also the formulae

$$\bigwedge_{i=1}^{n}(r_i \preccurlyeq s_i) \wedge \neg \bigvee_{j=1}^{m}(t_j \preccurlyeq u_j) \to (\bot \preccurlyeq \neg(r_i \preccurlyeq s_i))$$

and

$$\bigwedge_{i=1}^{n}(r_i \preccurlyeq s_i) \wedge \neg \bigvee_{j=1}^{m}(t_j \preccurlyeq u_j) \to (\bot \preccurlyeq (t_j \preccurlyeq u_j)) \ .$$

Thus by propositional reasoning we can derive the formula

$$\bigwedge_{i=1}^{n}(r_i \preccurlyeq s_i) \wedge \neg \bigvee_{j=1}^{m}(t_j \preccurlyeq u_j) \to \bigwedge_{i=1}^{n} (\bot \preccurlyeq \neg(r_i \preccurlyeq s_i)) \wedge \bigwedge_{j=1}^{m} (\bot \preccurlyeq (t_j \preccurlyeq u_j)) \ .$$

On the other hand it is not too hard to see that in $\mathsf{G}c\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$ we can derive for all formulae $A, B_1, \ldots, B_k$ the equivalence $\bigwedge_{i=1}^{k}(A \preccurlyeq B_i) \leftrightarrow (A \preccurlyeq \bigvee_{i=1}^{k} B_i)$, and thus we can derive the formula

$$\bigwedge_{i=1}^{n}(r_i \preccurlyeq s_i) \wedge \neg \bigvee_{j=1}^{m}(t_j \preccurlyeq u_j) \to \left(\bot \preccurlyeq \bigvee_{i=1}^{n} \neg(r_i \preccurlyeq s_i) \vee \bigvee_{j=1}^{m}(t_j \preccurlyeq u_u) \right)$$

which by standard propositional reasoning is equivalent to the formula $(\mathsf{A})_{i,j}$ in the ω-set for the axiom (A). Now by the translation of Chapter 3 the ω-set for the axiom (A) is translated into the rule $\mathsf{R_A}$. Thus using Theorem 3.2.30 the latter is equivalent over $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$ to the ω set for (A), and thus also to the two axioms (A1) and (A2). $\square$

Since all the additional rules contain at most one formula in their principal formulae, the results about the representations of the rules in the cut closure as small cut trees straightforwardly yield decidability and complexity results.

Theorem 5.3.2 (c)**.** *The validity problem for each of the logics $\mathbb{VN}_\preccurlyeq, \mathbb{VT}_\preccurlyeq, \mathbb{VW}_\preccurlyeq, \mathbb{VC}_\preccurlyeq$ is in* PSPACE*. The validity problem for the logics $\mathbb{VA}_\preccurlyeq$ and $\mathbb{VNA}_\preccurlyeq$ is in* EXPTIME*.*

Proof. The result for the logics without absoluteness follows directly since by Corollary 5.2.4 the rule set $\mathsf{cc}\left(\mathcal{R}'_{\mathbb{V}_\preccurlyeq}\right)$ is contraction closed, and thus by Corollary 4.1.30 the cut closures of extensions of $\mathcal{R}'_{\mathbb{V}_\preccurlyeq}$ with rules from $\mathsf{R_N}, \mathsf{R_T}, \mathsf{R_W}, \mathsf{R_C}$ are contraction closed (and therefore saturated) as well. For the logics involving absoluteness Corollary 4.1.30 only yields contraction closure of the cut closure of the sets $\mathcal{R}'_{\mathbb{V}_\preccurlyeq} \cup \{\mathsf{R_A}\}$ resp. $\mathcal{R}'_{\mathbb{V}_\preccurlyeq} \cup \{\mathsf{R_N}, \mathsf{R_A}\}$. But in these cases since the only restrictions occurring in the rules are $\mathcal{C}_\emptyset$ and $\mathcal{C}_{\mathbb{VA}}$, it is straightforward to check that the cut closure is mixed- and context-cut closed as well. Now Lemma 4.1.7 and Corollary 4.1.21 resp. Corollary 4.1.20 yield the complexity bounds. $\square$

For the logics $\mathbb{V}_\preccurlyeq, \mathbb{VN}_\preccurlyeq, \mathbb{VT}_\preccurlyeq, \mathbb{VW}_\preccurlyeq$ and $\mathbb{VC}_\preccurlyeq$ the PSPACE-complexity result given in the previous theorem resp. Corollary 5.2.5 is optimal. This can be seen by reduction from the validity problem for the standard modal logics K, D or T, which are embedded into the logic at hand using the translation $\Diamond p \leftrightarrow (p \preccurlyeq \top)$, see [Lew73a, Chapter 6].

While Theorem 5.3.2 gives us decision procedures and complexity bounds for the extensions of $\mathbb{V}_\preccurlyeq$ without having to explicitly construct cut-free sequent calculi, it is now not too difficult to do so by simply computing cuts between rules in $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$ and the additional rules. Since a cut between a rule $\mathsf{R}_{m,n}$ and the rule $\mathsf{R_N}$ gives the rule $\mathsf{R}_{m,0}$ in case $n = 1$ and is subsumed by the rule $\mathsf{R}_{m,n-1}$ in case $n > 1$, for the logic $\mathbb{VN}_\preccurlyeq$ it suffices to add all the rules $\mathsf{R}_{m,0}$ for $m \geq 1$ to the rule set. A cut between a rule $\mathsf{R}_{m,n}$ and the rule $\mathsf{R_T}$ on the other hand deletes all the heads pointing to the first argument of one positive literal of $\mathsf{R}_{m,n}$ and replaces the tails emerging from its second argument by the context. Thus in case $n > 1$ this rule is again subsumed by the rule $\mathsf{R}_{m,n-1}$. In case $n = 1$ we need to add the constructed rule T_m to the rule set. For $\mathbb{VW}_\preccurlyeq$ and $\mathbb{VC}_\preccurlyeq$ observe that a cut between the rule $\mathsf{R_W}$ resp. $\mathsf{R_C}$ and the rule $\mathsf{R}_{1,1}$ yields the rules

$$\mathsf{R_{W2}} := \{(\ \Rightarrow p; \mathcal{C}_{\mathsf{id}})\}/\ \Rightarrow (p \preccurlyeq q)$$
$$\mathsf{R_{C2}} := \{(p \Rightarrow\ ; \mathcal{C}_{\mathsf{id}}), (\ \Rightarrow q; \mathcal{C}_{\mathsf{id}})\}/(p \preccurlyeq q) \Rightarrow$$

which subsume the rules $\mathsf{R_W}$ and $\mathsf{R_C}$ respectively. Cutting $\mathsf{R_{W2}}$ with $\mathsf{R}_{m,n}$ has the effect of replacing the antecedent of one negative literal with the context in all premisses and deleting

$$\Sigma_m \Rightarrow \Pi_n := (p_1 \preccurlyeq q_1), \dots, (p_m \preccurlyeq q_m) \Rightarrow (r_1 \preccurlyeq s_1), \dots, (r_n \preccurlyeq s_n)$$

$$\begin{aligned}
\mathsf{T}_m &:= \mathcal{P}^{\mathsf{T}}_m / \Sigma_m \Rightarrow \Pi_0 \text{ where} \\
&\quad \mathcal{P}^{\mathsf{T}}_m := \begin{array}{l} \{(p_k \Rightarrow q_1, \dots, q_{k-1}; \mathcal{C}_\emptyset) \mid 1 \le k \le m\} \\ \cup \{(\Rightarrow q_1, \dots, q_m; \mathcal{C}_{\mathsf{id}})\} \end{array} \\
\mathsf{W}_{m,n} &:= \mathcal{P}^{\mathsf{W}}_{m,n} / \Sigma_m \Rightarrow \Pi_n \text{ where} \\
&\quad \mathcal{P}^{W}_{m,n} := \begin{array}{l} \{(p_k \Rightarrow q_1, \dots, q_{k-1}, r_1, \dots, r_n; \mathcal{C}_\emptyset) \mid 1 \le k \le m\} \\ \cup \{(\Rightarrow q_1, \dots, q_m, r_1, \dots, r_m; \mathcal{C}_{\mathsf{id}})\} \end{array} \\
\mathsf{A}_{m,n} &:= \mathcal{P}^{\mathsf{A}}_{m,n} / \Sigma_m \Rightarrow \Pi_n \text{ where} \\
&\quad \mathcal{P}^{A}_{m,n} := \begin{array}{l} \{(s_k \Rightarrow q_1, \dots, q_m, r_1, \dots, r_n; \mathcal{C}_{\mathbb{VA}}) \mid 1 \le k \le n\} \\ \cup \{(p_k \Rightarrow q_1, \dots, q_{k-1}, r_1, \dots, r_n; \mathcal{C}_{\mathbb{VA}}) \mid 1 \le k \le m\} \end{array} \\
\mathsf{R}_{\mathsf{W2}} &:= \{(\Rightarrow p; \mathcal{C}_{\mathsf{id}})\} / \ \Rightarrow (p \preccurlyeq q) \\
\mathsf{R}_{\mathsf{C2}} &:= \{(p \Rightarrow\ ; \mathcal{C}_{\mathsf{id}}), (\Rightarrow q; \mathcal{C}_{\mathsf{id}})\} / (p \preccurlyeq q) \Rightarrow
\end{aligned}$$

$$\begin{array}{rlrl}
\mathcal{R}_{\mathbb{VN}_\preccurlyeq} &:= \{\mathsf{R}_{m,n} \mid m+n \ge 1\} & \mathcal{R}_{\mathbb{VC}_\preccurlyeq} &:= \mathcal{R}_{\mathbb{V}_\preccurlyeq} \cup \{\mathsf{R}_{\mathsf{W2}}, \mathsf{R}_{\mathsf{C2}}\} \\
\mathcal{R}_{\mathbb{VT}_\preccurlyeq} &:= \mathcal{R}_{\mathbb{V}_\preccurlyeq} \cup \{\mathsf{T}_m \mid m \ge 1\} & \mathcal{R}_{\mathbb{VA}_\preccurlyeq} &:= \{\mathsf{A}_{m,n} \mid m \ge 0,\ n \ge 1\} \\
\mathcal{R}_{\mathbb{VW}_\preccurlyeq} &:= \mathcal{R}_{\mathbb{V}_\preccurlyeq} \cup \{\mathsf{W}_{m,n} \mid m+n \ge 1\} & \mathcal{R}_{\mathbb{VNA}_\preccurlyeq} &:= \{\mathsf{A}_{m,n} \mid m+n \ge 1\}
\end{array}$$

Table 5.4: The rules and rule sets for extensions of $\mathbb{V}_\preccurlyeq$.

all doodles pointing to its succedent. For $\mathbb{VC}_\preccurlyeq$ note that a cut between R_{C2} and $\mathsf{R}_{0,2}$ yields the rule R_{W2}. For $\mathbb{VA}_\preccurlyeq$ similar to the case of $\mathbb{VW}_\preccurlyeq$ a cut between a rule $\mathsf{R}_{m,n}$ and the rule R_{A} on the uppermost negative literal of the former in effect deletes this literal from the conclusion and changes the context restriction in every premiss to $\mathcal{C}_{\mathbb{VA}}$. The new rules for $\mathbb{VT}_\preccurlyeq$, $\mathbb{VW}_\preccurlyeq$ and $\mathbb{VA}_\preccurlyeq$ in non-graphical notation and the resulting rule sets are given in the following definition.

Definition 5.3.3. For $m, n \in \mathbb{N}$ the rules $\mathsf{T}_m, \mathsf{W}_{m,n}, \mathsf{A}_{m,n}, \mathsf{R}_{\mathsf{W2}}, \mathsf{R}_{\mathsf{C2}}$ as well as the rule sets $\mathcal{R}_\mathcal{L}$ for $\mathcal{L} \in \{\mathbb{VN}_\preccurlyeq, \mathbb{VT}_\preccurlyeq, \mathbb{VW}_\preccurlyeq, \mathbb{VC}_\preccurlyeq, \mathbb{VA}_\preccurlyeq, \mathbb{VNA}_\preccurlyeq\}$ are defined as in Table 5.4.

Note that in the case of $\mathbb{VC}_\preccurlyeq$ we arrived in a systematic and purely syntactic way at the sequent system corresponding to the tableau system given in [Gen92]. Again all the rule sets are sound for the respective logics by construction, but as the next lemma shows this can also be seen directly.

Lemma 5.3.4 (c)**.** *Let $\mathcal{L}$ be a logic in $\{\mathbb{VN}_\preccurlyeq, \mathbb{VT}_\preccurlyeq, \mathbb{VW}_\preccurlyeq, \mathbb{VC}_\preccurlyeq, \mathbb{VA}_\preccurlyeq, \mathbb{VNA}_\preccurlyeq\}$. Then the sequent system given by $\mathsf{Gc}\mathcal{R}_\mathcal{L}\mathsf{Con}$ is sound for $\mathcal{L}$.*

Proof. Again we show soundness of the rule sets by showing that whenever all premisses of an application of a rule from the rule set are valid in the corresponding class of sphere models, then so is its conclusion. For the rules in $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$ see Theorem 5.2.2. The proofs for the additional rules are all similar to the case for $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$.

For $\mathcal{R}_{\mathbb{VN}_{\preccurlyeq}}$: Assume that for $m > 0$ the premisses of an application of the rule $\mathsf{R}_{m,0}$ are $\mathbb{VN}_{\preccurlyeq}$-valid. Then for every k with $1 \leq k \leq m$ we have $\models_{\mathbb{VN}_{\preccurlyeq}} A_k \to \bigvee_{\ell=1}^{k-1} B_\ell$ and moreover we have $\models_{\mathbb{VN}_{\preccurlyeq}} \bigvee_{\ell=1}^{m} B_\ell$. Suppose that $\mathcal{I}$ is a normal sphere model, i.e., for every world $i \in I$ we have $\bigcup \$_i \neq \emptyset$, and take an arbitrary world $i \in I$. We need to show that $i \in [\![\neg \bigwedge_{\ell=1}^{m}(A_\ell \preccurlyeq B_\ell)]\!]$. Assume on the contrary that $i \in [\![A_\ell \preccurlyeq B_\ell]\!]$ for every $\ell \leq m$ and take a nonempty sphere $S \in \$_i$. Since $\bigvee_{\ell=1}^{m} B_\ell$ is $\mathbb{VN}_{\preccurlyeq}$-valid, and since S is nonempty, there is a world $i_1 \in S$ with $i_1 \in \bigcup_{\ell=1}^{m} [\![B_\ell]\!]$ and thus we find an index k_1 such that $i_1 \in [\![B_{k_1}]\!]$. But since by assumption $i \in [\![A_{k_1} \preccurlyeq B_{k_1}]\!]$ we find a world $i_2 \in S$ such that $i_2 \in [\![A_{k_1}]\!]$. On the other hand we have $\models_{\mathbb{VN}_{\preccurlyeq}} A_{k_1} \to \bigvee_{\ell=1}^{k_1-1} B_\ell$ and thus we obtain $i_2 \in [\![B_{k_2}]\!]$ for an index k_2 with $k_2 < k_1$. Continuing like this after at most m steps we find a world j with $j \in S \cap [\![A_1]\!]$ in contradiction to the fact that $A_1 \to \bot$ is $\mathbb{VN}_{\preccurlyeq}$-valid. Thus we have $i \notin [\![A_\ell \preccurlyeq B_\ell]\!]$ for some $\ell \leq m$.

For $\mathcal{R}_{\mathbb{VT}_{\preccurlyeq}}$: Suppose that the premisses of an application of the rule T_m are $\mathbb{VT}_{\preccurlyeq}$-valid. Then again for every k with $1 \leq k \leq m$ we have $\models_{\mathbb{VT}_{\preccurlyeq}} A_k \to \bigvee_{\ell=1}^{k-1} B_\ell$ and furthermore $\models_{\mathbb{VT}_{\preccurlyeq}} \bigwedge \Gamma \to \bigvee \Delta \vee \bigvee_{\ell=1}^{m} B_\ell$. Now let $\mathcal{I}$ be totally reflexive, i.e., for every world $i \in I$ we have $i \in \bigcup \$_i$. Then for any world $i \in I$ we have $i \in [\![\bigwedge \Gamma \to \bigvee \Delta]\!]$ or $i \in [\![\bigvee_{\ell=1}^{m} B_\ell]\!]$. In the former case we also have $i \in [\![\bigwedge \Gamma \wedge \bigwedge_{\ell=1}^{m}(A_\ell \preccurlyeq B_\ell) \to \bigvee \Delta]\!]$ and are done. In the latter case since $\mathcal{I}$ is totally reflexive we can choose a sphere $S \in \$_i$ with $i \in S$. Now the same proof as in the case for $\mathbb{VN}_{\preccurlyeq}$ above shows that for some $\ell \leq m$ we have $i \notin [\![A_\ell \preccurlyeq B_\ell]\!]$.

For $\mathcal{R}_{\mathbb{VW}_{\preccurlyeq}}$: If for some $m, n \geq 0$ the premisses of an application of the rule $\mathsf{W}_{m,n}$ are valid, we have

$$\models_{\mathbb{VW}_{\preccurlyeq}} \bigwedge \Gamma \to \bigvee_{k=1}^{m} B_k \vee \bigvee_{k=1}^{n} C_k \vee \bigvee \Delta$$

and furthermore for every k with $1 \leq k \leq m$ we have

$$\models_{\mathbb{VW}_{\preccurlyeq}} A_k \to \bigvee_{\ell=1}^{k-1} B_\ell \vee \bigvee_{\ell=1}^{n} C_\ell \; .$$

Now suppose that the sphere model $\mathcal{I}$ is weakly centered, i.e., for every world $i \in I$ there is a sphere $S \in \$_i$ with $S \neq \emptyset$ and for every sphere $S \in \$_i$ with $S \neq \emptyset$ we have $i \in S$. Now take an arbitrary world $i \in I$. Again we need to show that

$$i \in \left[\!\!\left[\bigwedge \Gamma \wedge \bigwedge_{k=1}^{m} (A_k \preccurlyeq B_k) \to \bigvee_{k=1}^{n} (C_k \preccurlyeq D_k) \vee \bigvee \Delta \right]\!\!\right] \; .$$

By assumption we have $i \in [\![\bigwedge \Gamma \to \bigvee_{k=1}^{m} B_k \vee \bigvee_{k=1}^{n} C_k \vee \bigvee \Delta]\!]$. Then we have either $i \notin \bigcup_{\ell=1}^{m} [\![B_\ell]\!] \cup \bigcup_{\ell=1}^{n} [\![C_\ell]\!]$ and are done since then $i \in [\![\bigwedge \Gamma \to \bigvee \Delta]\!]$; or we have $i \in [\![C_\ell]\!]$ for some ℓ with $1 \leq \ell \leq n$ and are done since the world i is contained in every non-empty sphere $S \in \$_i$ and hence $i \in [\![C_k \preccurlyeq D_k]\!]$; or we have $i \in [\![B_k]\!]$ for some k with $1 \leq k \leq m$. In the latter case we fix such a k and use the fact that $\mathcal{I}$ is weakly centered to choose a non-empty sphere $S \in \$_i$. Then $i \in S$. Again, for the sake of contradiction assume that $i \notin [\![\bigwedge_{k=1}^m (A_k \preccurlyeq B_k) \to \bigvee_{k=1}^n (C_k \preccurlyeq D_k)]\!]$. Then in particular for every $k \leq m$ we have $i \in [\![A_k \preccurlyeq B_k]\!]$ and for every $k \leq n$ we have $i \notin [\![C_k \preccurlyeq D_k]\!]$. Then since $i \in [\![B_k]\!]$ there is a world $i_1 \in S \cap [\![A_k]\!]$. Since $\models_{\mathbb{VW}_\preccurlyeq} A_k \to \bigvee_{\ell=1}^{k-1} B_\ell \vee \bigvee_{\ell=1}^n C_\ell$ we have $i_1 \in \bigcup_{\ell=1}^{k-1} [\![B_\ell]\!] \cup \bigcup_{\ell=1}^n [\![C_\ell]\!]$. Repeating the argument at most k times yields a world $j \in S \cap \bigcup_{\ell=1}^n [\![C_\ell]\!]$. Choose k_1 with $j \in [\![C_{k_1}]\!]$. Since by assumption $i \notin [\![C_{k_1} \preccurlyeq D_{k_1}]\!]$, there is a sphere $S' \subsetneqq S$ with $S' \cap [\![C_{k_1}]\!] = \emptyset$ and $S' \cap [\![D_{k_1}]\!] \neq \emptyset$. Since $\mathcal{I}$ is weakly centered again we have $i \in S'$. Now repeating the whole argument with the sphere S' instead of S gives a sphere $S'' \subsetneqq S'$ with $S'' \cap ([\![C_{k_1}]\!] \cup [\![C_{k_2}]\!]) = \emptyset$. Continuing like this after at most n repetitions we arrive at a sphere S''' with $S \cap [\![C_k]\!] = \emptyset$ for every $k \leq n$. But now a final iteration of the first part of the argument yields $S''' \cap \bigcup_{\ell=1}^n [\![C_\ell]\!] \neq \emptyset$, a contradiction.

For $\mathcal{R}_{\mathbb{VC}_\preccurlyeq}$: Since every centered sphere model is also normal, and since the rule $\mathsf{R_{W2}}$ is the same as the rule $\mathsf{W}_{1,0}$, we have soundness of the rule $\mathsf{R_{W2}}$ as above. For rule $\mathsf{R_{C2}}$ suppose on the other hand that $\models_{\mathbb{VC}_\preccurlyeq} \bigwedge \Gamma \wedge A \to \bigvee \Delta$ and $\models_{\mathbb{VC}_\preccurlyeq} \bigwedge \Gamma \to B \vee \bigvee \Delta$. Then since $\mathcal{I}$ is a centered sphere model for all $i \in I$ we have $\{i\} \in \$_i$, and for $i \in I$ obviously $\{i\}$ is the smallest sphere in $\$_i$. We need to show that $i \in [\![\bigwedge \Gamma \wedge (A \preccurlyeq B) \to \Delta]\!]$. If $i \in [\![\bigwedge \Gamma \Rightarrow \bigvee \Delta]\!]$, then $i \in [\![\bigwedge \Gamma \wedge (A \preccurlyeq B) \to \bigvee \Delta]\!]$ as well and we are done. Otherwise since $\models_{\mathbb{VC}_\preccurlyeq} \bigwedge \Gamma \wedge A \Rightarrow \bigvee \Delta$ and $\models_{\mathbb{VC}_\preccurlyeq} \bigwedge \Gamma \to B \vee \bigvee \Delta$ we have $i \notin [\![A]\!]$ and $i \in [\![B]\!]$. But then since $\{i\}$ is the smallest sphere in $\$_i$ we have $i \notin [\![A \preccurlyeq B]\!]$ and thus $i \in [\![\bigwedge \Gamma \wedge (A \preccurlyeq B) \to \bigvee \Delta]\!]$.

For $\mathcal{R}_{\mathbb{VA}_\preccurlyeq}$: This case again is analogous to the case of $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$, making use of the easy fact that if in a sphere model for two worlds i, j we have $\$_i = \$_j$, then for every formula $(A \preccurlyeq B)$ we have $i \in [\![A \preccurlyeq B]\!]$ if and only if $j \in [\![A \preccurlyeq B]\!]$. In more detail, assume that the premisses of an application of $\mathsf{A}_{m,n}$ for $m \geq 0$ and $n \geq 1$ are valid, i.e., suppose that for every k with $1 \leq k \leq n$ we have

$$\models_{\mathbb{VA}_\preccurlyeq} \bigwedge_{\ell=1}^{c} (E_\ell \preccurlyeq F_\ell) \wedge D_k \to \bigvee_{\ell=1}^{m} B_\ell \vee \bigvee_{\ell=1}^{n} C_\ell \vee \bigvee_{\ell=1}^{d} (G_\ell \preccurlyeq H_\ell)$$

and that for every k with $1 \leq k \leq m$ we have

$$\models_{\mathbb{VA}_\preccurlyeq} \bigwedge_{\ell=1}^{c} (E_\ell \preccurlyeq F_\ell) \wedge A_k \to \bigvee_{\ell=1}^{k-1} B_\ell \vee \bigvee_{\ell=1}^{n} C_\ell \vee \bigvee_{\ell=1}^{d} (G_\ell \preccurlyeq H_\ell) \, .$$

Furthermore, let $\mathcal{I}$ be an absolute sphere model, i.e., for every two worlds i, j from $\mathcal{I}$ we have $\$_i = \$_j$. Take an arbitrary world i from $\mathcal{I}$. We need to show that

$$i \in \left[\!\!\left[\bigwedge_{\ell=1}^{c} (E_\ell \preccurlyeq F_\ell) \wedge \bigwedge_{\ell=1}^{m} (A_\ell \preccurlyeq B_\ell) \to \bigvee_{\ell=1}^{n} (C_\ell \preccurlyeq D_\ell) \vee \bigvee_{\ell=1}^{d} (G_\ell \preccurlyeq H_\ell) \right]\!\!\right] \, .$$

So assume that $i \in [\![\bigwedge_{\ell=1}^c (E_\ell \preccurlyeq F_\ell) \wedge \bigwedge_{\ell=1}^m (A_\ell \preccurlyeq B_\ell)]\!]$, that $i \notin [\![G_\ell \preccurlyeq H_\ell]\!]$ for every ℓ with $1 \leq \ell \leq d$ and that for a k with $1 \leq k \leq n$ we have $i \notin [\![C_\ell \preccurlyeq D_\ell]\!]$ for every ℓ with

$1 \leq \ell \leq n$ and $\ell \neq k$. Then if $\bigcup \$_i \cap [\![D_k]\!] = \emptyset$ we are done. In particular this holds in case $\bigcup \$_i = \emptyset$. Otherwise, choose a world $j \in \bigcup \$_i \cap [\![D_k]\!]$. Using validity of $\bigwedge_{\ell=1}^{c}(E_i \preccurlyeq F_i) \wedge D_k \to \bigvee_{\ell=1}^{m} B_\ell \vee \bigvee_{\ell=1}^{n} C_\ell \vee \bigvee_{\ell=1}^{d}(G_\ell \preccurlyeq H_\ell)$ we have

$$j \in \left[\!\!\left[\bigwedge_{\ell=1}^{c}(E_\ell \preccurlyeq F_\ell) \to \bigvee_{\ell=1}^{m} B_\ell \vee \bigvee_{\ell=1}^{n} C_\ell \vee \bigvee_{\ell=1}^{d}(G_\ell \preccurlyeq H_\ell) \right]\!\!\right] .$$

But since $\mathcal{I}$ is absolute the same formulae of the form $X \preccurlyeq Y$ hold in the worlds i and j, and thus we have $j \in [\![\bigwedge_{\ell=1}^{c}(E_\ell \preccurlyeq F_\ell)]\!]$ and $j \notin [\![G_\ell \preccurlyeq H_\ell]\!]$ for every ℓ with $1 \leq \ell \leq d$. Hence we obtain that $j \in \bigcup_{\ell=1}^{m} [\![B_\ell]\!] \cup \bigcup_{\ell=1}^{n} [\![C_\ell]\!]$. If $j \in [\![B_\ell]\!]$ for some $\ell \leq m$, then since $i \in [\![A_\ell \preccurlyeq B_\ell]\!]$ we find a world $j_2 \in S \cap [\![A_\ell]\!]$. Thus using validity of $\bigwedge_{\ell'=1}^{c}(E'_\ell \preccurlyeq F'_\ell) \wedge A_\ell \to \bigvee_{\ell'=1}^{\ell-1} B_{\ell'} \vee \bigvee_{\ell'=1}^{n} C_{\ell'} \vee \bigvee_{\ell'=1}^{d}(G_{\ell'} \preccurlyeq H_{\ell'})$ and the fact that the same formulae of the form $X \preccurlyeq Y$ hold in the worlds i and j_2 we get that $j_2 \in \bigcup_{\ell'=1}^{\ell-1} [\![B_{\ell'}]\!] \cup \bigcup_{\ell'=1}^{n} [\![C_{\ell'}]\!]$. Iterating this process we find a world j' with $j' \in S \cap \bigcup_{\ell=1}^{n} [\![C_\ell]\!]$. If $j' \notin [\![C_k]\!]$, then there is an index ℓ with $\ell \neq k$ with $j' \in [\![C_\ell]\!]$. On the other hand we have $i \notin [\![C_\ell \preccurlyeq D_\ell]\!]$, and thus there is a sphere $S' \in \$_i$ with $S' \subsetneqq S$ and $S' \cap [\![D_\ell]\!] \neq \emptyset$ and $S' \cap [\![C_\ell]\!] = \emptyset$. Similar to above there is a world $j'' \in S' \cap \bigcup_{t=1}^{n} [\![C_t]\!] = (S' \cap \bigcup_{t=1}^{n} [\![C_t]\!]) \setminus [\![C_\ell]\!]$. Iterating the argument we find a sphere $S'' \in \$_i$ with $\emptyset \neq S'' \cap \bigcup_{\ell=1}^{n} [\![C_\ell]\!] = S'' \cap [\![C_k]\!]$ and $S'' \subseteq S$. Thus we get $i \in [\![C_k \preccurlyeq D_k]\!]$ and we are done.

For $\mathcal{R}_{\mathbb{VNA}_\preccurlyeq}$: For an application of the rule $\mathsf{A}_{m,n}$ with $n \geq 1$ this follows as in the case for $\mathbb{VA}_\preccurlyeq$. For the case $n = 0$ it is easy to adapt the proof for $\mathbb{VN}_\preccurlyeq$ to the additional context in the fashion of the previous case. □

The next step is to show cut-free completeness. We know from Proposition 5.3.1 that the extensions of the calculus $\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{CutCon}$ with combinations of the rules corresponding to the axioms characterising the logics are complete for the respective logics. Thus taking the logic $\mathbb{VC}_\preccurlyeq$ as an example we have that since $\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R_C}\mathsf{CutCon}$ is complete for $\mathbb{VC}_\preccurlyeq$, by cut elimination the system $\mathsf{Gc}\,\mathsf{cc}\left(\mathcal{R}_{\mathbb{V}_\preccurlyeq} \cup \{\mathsf{R_C}\}\right)\mathsf{Con}$ is complete for $\mathbb{VC}_\preccurlyeq$ as well. Now if we can show that all rules in $\mathsf{cc}\left(\mathcal{R}_{\mathbb{V}_\preccurlyeq} \cup \{\mathsf{R_C}\}\right)$ are derivable rules in $\mathcal{R}_{\mathbb{VC}_\preccurlyeq}\mathsf{ConW}$, then we are done, since then as noticed in Remark 2.4.18 we can transform every (cut-free) derivation in $\mathsf{Gc}\,\mathsf{cc}\left(\mathcal{R}_{\mathbb{V}_\preccurlyeq} \cup \{\mathsf{R_C}\}\right)\mathsf{Con}$ into a (cut-free) derivation in $\mathsf{Gc}\mathcal{R}_{\mathbb{VC}_\preccurlyeq}\mathsf{Con}$. As it might be helpful to visualise the rules, the rule doodles representing the rules $\mathsf{R_N}, \mathsf{R_T}$ and $\mathsf{R_W}$ are given in Figure 5.4.

Lemma 5.3.5 (c)**.** *Let $\mathcal{L}$ be a logic in $\{\mathbb{VN}_\preccurlyeq, \mathbb{VT}_\preccurlyeq, \mathbb{VW}_\preccurlyeq, \mathbb{VC}_\preccurlyeq, \mathbb{VA}_\preccurlyeq, \mathbb{VNA}_\preccurlyeq\}$. Then the sequent system given by $\mathsf{Gc}\mathcal{R}_\mathcal{L}\mathsf{Con}$ is complete for $\mathcal{L}$.*

Proof. We show that if $\mathcal{L}$ is one of the logics considered and $\mathcal{R}$ is the rule set for this logic given in Proposition 5.3.1, then all rules in $\mathsf{cc}\,(\mathcal{R})$ are derivable rules in $\mathcal{R}_\mathcal{L}\mathsf{WCon}$. In order to show this it is enough to show that cuts between a rule $\mathsf{R}_{m,n}$ and possibly several instances of the additional rules not in $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$ are derivable rules in $\mathcal{R}_\mathcal{L}$. This is due to the fact that as we

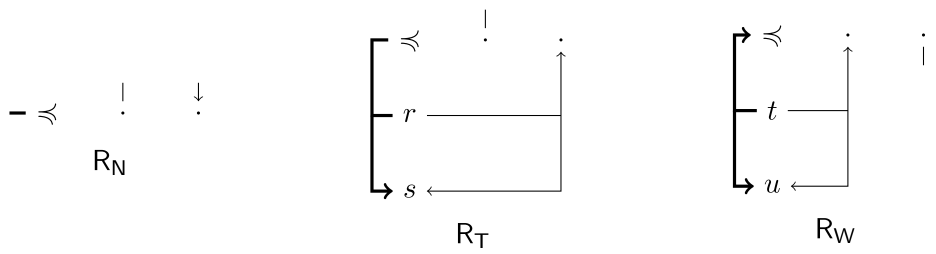


Figure 5.4: The graphical representations of the rules for extensions of $\mathbb{V}_{\preccurlyeq}$

have seen in Theorem 5.2.3 cuts between two rules $\mathsf{R}_{m,n}$ and $\mathsf{R}_{k,\ell}$ are subsumed by the single rule $\mathsf{R}_{m+k-1,n+\ell-1}$, and thus if the construction of a rule in $\mathsf{cc}(\mathcal{R})$ involves such a cut, then reordering the cut tree and replacing this cut with the rule $\mathsf{R}_{m+k-1,n+\ell-1}$ gives a rule without the cut which subsumes the original rule. We consider each of the rule sets in turn.

For $\mathcal{R}_{\mathbb{VN}_{\preccurlyeq}}$: A cut between the rule $\mathsf{R_N} = \{(p \Rightarrow ;\mathcal{C}_\emptyset), (\Rightarrow q;\mathcal{C}_\emptyset)\}/(p \preccurlyeq q) \Rightarrow$ and a rule $\mathsf{R}_{m,n}$ with $n \geq 1$ simply deletes one positive literal in the conclusion of the latter, all heads pointing to its first argument and all tails emerging from its second argument in the premisses. Thus the resulting rule is subsumed by the rule $\mathsf{R}_{m,n-1}$.

For $\mathcal{R}_{\mathbb{VT}_{\preccurlyeq}}$: A cut between the rule $\mathsf{R_T} = \{(p \Rightarrow \ \mathcal{C}_\emptyset), (\Rightarrow q;\mathcal{C}_{\mathsf{id}})\}/(p \preccurlyeq q) \Rightarrow$ and a rule $\mathsf{R}_{m,n}$ with $n \geq 2$ similarly deletes a positive literal in the conclusion of the latter and all heads pointing to its first argument. This is already enough to ensure that the resulting rule is subsumed by the rule $\mathsf{R}_{m,n-1}$. If $n = 1$, then the cut gives the rule T_m.

For $\mathcal{R}_{\mathbb{VW}_{\preccurlyeq}}$: The rules $\mathsf{R_T}$ and $\mathsf{R_W}$ are subsumed by the rules $\mathsf{W}_{1,0}$ resp. $\mathsf{W}_{0,1}$. Cuts between the rules $\mathsf{R_T}$ and $\mathsf{R}_{m,n}$ are subsumed by the rules $\mathsf{R}_{m,n-1}$ for $n > 1$ or $\mathsf{W}_{m,0}$ for $n = 1$. Thus given cut between a rule $\mathsf{R}_{m,n}$ and multiple instances of the rules $\mathsf{R_T}$ and $\mathsf{R_W}$ it suffices to first eliminate the cuts involving the instance of $\mathsf{R_T}$ and then deal with the remaining cuts between a rule $\mathsf{R}_{m,n}$ or $\mathsf{W}_{m,0}$ and instances of $\mathsf{R_W}$. Cuts between a rule $\mathsf{R}_{m,n}$ and multiple instances of the rule $\mathsf{R_W} = \{(\Rightarrow p;\mathcal{C}_{\mathsf{id}}), (q \Rightarrow ;\mathcal{C}_\emptyset)\}/ \Rightarrow (p \preccurlyeq q)$ are treated in the following way. Assume that the cuts occur on the (negative) literals $\{(p_i \preccurlyeq q_i) \mid i \in I\}$ of $\mathsf{R}_{m,n}$ for some $I \subseteq \{1, \dots, m\}$. Then the cuts with $\mathsf{R_W}$ have the effect of deleting all heads pointing to the second arguments of the corresponding literals. Moreover, the cut on the literal $(p_{\max(I)} \preccurlyeq q_{\max(I)})$ produces a premiss which can be weakened to the premiss with heads pointing to all q_i for $i \in \{1, \dots, m\} \smallsetminus I$ and to all r_j for $1 \leq j \leq n$ and with heads pointing to resp. tails emerging from all the context formulae. Thus the resulting rule is subsumed by the rule $\mathsf{W}_{m-|I|,n}$. In the case of a cut between the rules $\mathsf{W}_{m,0}$ and $\mathsf{R_W}$ the resulting rule is subsumed by the rule $\mathsf{W}_{m-1,0}$.

For $\mathcal{R}_{\mathbb{VC}_{\preccurlyeq}}$: The rule $\mathsf{R_C}$ is subsumed by the rule $\mathsf{R_{C2}}$. On the other hand, a cut between the rule $\mathsf{R}_{m,n+\ell}$ and ℓ instances of the rule $\mathsf{R_C}$ on the literals $(r_{n+1} \preccurlyeq s_{n+1}), \dots, (r_{n+\ell} \preccurlyeq s_{n+\ell})$ has the premisses

$$\begin{aligned}\mathcal{P}'_{m,n} = &\{(s_k \Rightarrow q_1, \dots, q_m, r_1, \dots, r_n; \mathcal{C}_{\mathsf{id}}) \mid 1 \leq k \leq n\}\\ &\cup \{(p_k \Rightarrow q_1, \dots, q_{k-1}, r_1, \dots, r_n; \mathcal{C}_{\mathsf{id}}) \mid 1 \leq k \leq m\}\\ &\cup \{(\ \Rightarrow q_1, \dots, q_m, r_1, \dots, r_n; \mathcal{C}_{\mathsf{id}})\} \ .\end{aligned}$$

Thus an application of this rule is derived using multiple applications of first the rule R_{W2} and then the rule R_{C2} as follows. First we have a derivation $\mathcal{D}$ given by

$$\frac{\Gamma \Rightarrow q_1, \dots, q_m, r_1, \dots, r_n, \Delta}{\Gamma \Rightarrow q_1, \dots, q_m, (r_1 \preccurlyeq s_1), \dots, (r_n \preccurlyeq s_n), \Delta}\ \mathsf{R}_{\mathsf{W2}}$$

and for k with $1 \leq k \leq m$ derivations $\mathcal{D}_k$ given by

$$\frac{\dfrac{\Gamma, p_k \Rightarrow q_1, \dots, q_{k-1}, r_1, \dots, r_n, \Delta}{\Gamma, p_k \Rightarrow q_1, \dots, q_{k-1}, (r_1 \preccurlyeq s_1), \dots, (r_n \preccurlyeq s_n), \Delta}\ \mathsf{R}_{\mathsf{W2}}}{\Gamma, p_k, (p_{k+1} \preccurlyeq q_{k+1}), \dots, (p_m \preccurlyeq q_m) \Rightarrow q_1, \dots, q_{k-1}, (r_1 \preccurlyeq s_1), \dots, (r_n \preccurlyeq s_n), \Delta}\ \mathsf{W}$$

Abbreviating the multiset $(r_1 \preccurlyeq s_1), \dots, (r_n \preccurlyeq s_n), \Delta$ to Σ we then derive the conclusion using multiple applications of R_{C2}:

$$\frac{\begin{matrix}\mathcal{D}_1\\ \vdots \\ \Gamma, p_1, (p_2 \preccurlyeq q_2), \dots, (p_m \preccurlyeq q_m) \Rightarrow \Sigma\end{matrix} \qquad \begin{matrix}\dfrac{\begin{matrix}\mathcal{D}_m\\ \vdots\\ \Gamma, p_m \Rightarrow q_1, \dots, q_{m-1}, \Sigma\end{matrix} \qquad \begin{matrix}\mathcal{D}\\ \vdots\\ \Gamma \Rightarrow q_1, \dots, q_m, \Sigma\end{matrix}}{\Gamma, (p_m \preccurlyeq q_m) \Rightarrow q_1, \dots, q_{m-2}, \Sigma}\ \mathsf{R}_{\mathsf{C2}}\\ \vdots\\ \Gamma, (p_2 \preccurlyeq q_2), \dots, (p_m \preccurlyeq q_m) \Rightarrow q_1 \Sigma\end{matrix}}{\Gamma, (p_1 \preccurlyeq q_1), \dots, (p_m \preccurlyeq q_m) \Rightarrow \Sigma}\ \mathsf{R}_{\mathsf{C2}}$$

This shows that the original rule is a derivable rule in $\mathcal{R}_{\mathbb{VC}_{\preccurlyeq}}\mathsf{ConW}$.

For $\mathcal{R}_{\mathbb{VA}_{\preccurlyeq}}$: The rule R_{A} is subsumed by the rule $\mathsf{A}_{0,1}$, and for $m \geq 0, n \geq 1$ the rule $\mathsf{R}_{m,n}$ is subsumed by the rule $\mathsf{A}_{m,n}$. A cut between a rule $\mathsf{R}_{m,n}$ and the rule R_{A} on the literal $(p_i \preccurlyeq q_i)$ deletes all doodles representing premisses of $\mathsf{R}_{m,n}$ with tails emerging from p_i and replaces all heads in doodles representing the premisses of $\mathsf{R}_{m,n}$ pointing to q_i with heads (resp. tails) pointing to (resp. emerging from) context formulae $(t \preccurlyeq u)$ and $(v \preccurlyeq w)$. Thus cuts between $\mathsf{R}_{m,n}$ and ℓ instances of R_{A} are subsumed by the rule $\mathsf{A}_{m-\ell,n}$.

For $\mathcal{R}_{\mathbb{VNA}_{\preccurlyeq}}$: Again the rules R_{A} and $\mathsf{R}_{m,n}$ are subsumed by the rules $\mathsf{A}_{0,1}$ resp. $\mathsf{A}_{m,n}$. The rule R_{N} furthermore is subsumed by the rule $\mathsf{A}_{1,0}$. A cut between R_{N} and R_{A} gives the identity rule. Cuts between a rule $\mathsf{R}_{m,n}$ and k instances of the rule R_{A} and ℓ instances of the rule R_{N} are subsumed by the rule $\mathsf{A}_{m-k,n-\ell}$. $\square$

Together, these two lemmata give soundness and cut-free completeness of our rule sets.

Theorem 5.3.6 (c)**.** *Let $\mathcal{L}$ be a logic in $\{\mathbb{VN}_{\preccurlyeq}, \mathbb{VT}_{\preccurlyeq}, \mathbb{VW}_{\preccurlyeq}, \mathbb{VC}_{\preccurlyeq}, \mathbb{VA}_{\preccurlyeq}, \mathbb{VNA}_{\preccurlyeq}\}$. Then the sequent system given by $\mathsf{Gc}\mathcal{R}_{\mathcal{L}}\mathsf{Con}$ is sound and complete for $\mathcal{L}$.*

Proof. Immediate from Lemma 5.3.4 and Lemma 5.3.5. □

Remark 5.3.7. For the logics $\mathbb{VN}_{\preccurlyeq}, \mathbb{VW}_{\preccurlyeq}, \mathbb{VA}_{\preccurlyeq}$ and $\mathbb{VNA}_{\preccurlyeq}$ the proof of Lemma 5.3.5 is also easily extended to show that the corresponding rule sets are principal-cut closed. Thus for these logics we might also use derivability of the original rules and the generic cut elimination theorem to show cut-free completeness. For the logics $\mathbb{VT}_{\preccurlyeq}$ and $\mathbb{VC}_{\preccurlyeq}$ we would need to add further rules to the rule set, though, which is why this presentation was chosen.

We may now also use the explicitly constructed rule sets modified according to Definition 2.6.1 to show decidability of the logics in question.

Corollary 5.3.8 (c)**.** *The validity problem for the logics $\mathbb{VN}_{\preccurlyeq}, \mathbb{VT}_{\preccurlyeq}, \mathbb{VW}_{\preccurlyeq}, \mathbb{VC}_{\preccurlyeq}$, is solvable in space polynomial in the circuit size of the input. The validity problem for $\mathbb{VA}_{\preccurlyeq}$ and $\mathbb{VNA}_{\preccurlyeq}$ is solvable in time exponential in the circuit size of the input.*

Proof. Let $\mathcal{L}$ be one of the specified logics. Then the rule set $\mathcal{R}_{\mathcal{L}}$ is tractable. Moreover, by Theorem 5.3.6 above and Theorem 2.6.5 the sequent calculus $(\mathsf{Gc}\mathcal{R}_{\mathcal{L}})^*$ is sound and complete for $\mathcal{L}$. Now using the procedures given in the proofs of Theorem 2.7.8 resp. Theorem 2.7.5 we obtain the complexity bound. □

5.4 Strong And Weak Counterfactual Implication

One of the original motivations for studying conditional logics stems from trying to formalise counterfactual implications. But from this point of view it is more natural to take either the strong counterfactual implication $\Box\!\!\Rightarrow$ or its weak version $\Box\!\!\rightarrow$ as a primitive instead of the comparative plausibility operator $\preccurlyeq$. Of course since these connectives are interdefinable a decision procedure for a conditional logic formulated in terms of one of these connectives immediately yields a decision procedure for the logic formulated in terms of the other ones. It is not immediately clear, however, that the resulting procedure has the same complexity as the original one. It is clear from the earlier mentioned translations of the strong and weak counterfactual implications into the comparative plausibility operator given by

$$(A \Box\!\!\Rightarrow B) \leftrightarrow \neg((A \wedge \neg B) \preccurlyeq (A \wedge B))$$
$$(A \Box\!\!\rightarrow B) \leftrightarrow (\bot \preccurlyeq A) \vee \neg((A \wedge \neg B) \preccurlyeq (A \wedge B))$$

that the (formula) size might grow exponentially. On the other hand, the number of subformulae grows only by a constant factor. Thus using the circuit representation of formulae (Definition 2.7.2) we do obtain complexity results for these languages.

Corollary 5.4.1 (c). *Let $\mathcal{L}$ be one of $\{\mathbb{V}, \mathbb{VN}, \mathbb{VT}, \mathbb{VW}, \mathbb{VC}\}$ and let $* \in \{\Box\!\!\Rightarrow, \Box\!\!\rightarrow\}$. Then the validity problem for the logic $\mathcal{L}_*$ is in* Pspace. *If $\mathcal{L} \in \{\mathbb{VA}, \mathbb{VNA}\}$, then the validity problem for $\mathcal{L}_*$ is in* Exptime.

Proof. Translating a formula A of $\mathcal{F}(\Box\!\!\Rightarrow)$ given in the circuit representation into the language of $\preccurlyeq$ adds 3 additional nodes for each node labelled with $\Box\!\!\Rightarrow$. Similarly, the translation from the language of $\Box\!\!\rightarrow$ adds 6 additional nodes for each node labelled with $\Box\!\!\rightarrow$. Thus in either case if A^τ is the translated formula we have $\|A^\tau\|_c \leq 6 \cdot \|A\|_c$ and are done using Corollary 5.2.5 resp. Corollary 5.3.8. $\square$

This reproves the known complexity results for the logics not including absoluteness from [FH94] in a purely syntactical way. For the logics $\mathbb{VA}_{\Box\!\rightarrow}$ and $\mathbb{VNA}_{\Box\!\rightarrow}$ our Exptime-complexity bound unfortunately is far from the coNP-bound established in the same article, but on the other hand this might be expected of a generic procedure. Even though we have decision procedures already, it is still interesting to construct explicit sequent calculi for the logics considered in the language of counterfactual implication. We can do so for the strong counterfactual implication by basically translating the rules for the comparative plausibility operator into the other language. The idea for this is to make use of the translation axiom

$$(A \Box\!\!\Rightarrow B) \leftrightarrow \neg (A \wedge \neg B \preccurlyeq A \wedge B) \tag{5.1}$$

from Section 5.1 to construct two "translation rules". The extension of the rule set $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$ with these translation rules then gives a rule set for the logic $\mathbb{V}_{\preccurlyeq,\Box\!\Rightarrow}$. Moreover, since both for the strong counterfactual implication and the comparative plausibility operator the principal formulae of the translation rules contain exactly one literal with this connective as its main connective, computing a cut between such a translation rule and a rule $\mathsf{R}_{m,n}$ has the effect of replacing one literal in the principal formulae of the latter rule with a literal with the strong counterfactual implication as main connective. Thus computing all possible cuts between a rule $\mathsf{R}_{m,n}$ and the two translation rules yields a rule corresponding to $\mathsf{R}_{m,n}$ but in the language of $\Box\!\!\Rightarrow$. The set of all these rules will then yield a sequent calculus for $\mathbb{V}_{\Box\!\Rightarrow}$. For extensions of $\mathbb{V}_{\Box\!\Rightarrow}$ the process is similar. The reason why this method does not work straightforwardly for the weak counterfactual implication is that the principal formulae of one of the translation rules constructed from the translation

$$(A \Box\!\!\rightarrow B) \leftrightarrow (\bot \preccurlyeq A) \vee \neg(A \wedge \neg B \preccurlyeq A \wedge B)$$

contain *two* literals with the comparative plausibility operator as main connective. Thus a cut between this rule and a rule $\mathsf{R}_{m,n}$ does not reduce the number of literals with the comparative plausibility operator as main connective in the principal formulae of the latter rule. On the other hand, we could try to use the translation

$$(A \preccurlyeq B) \leftrightarrow \big(A \vee B \mathrel{\Box\!\!\rightarrow} \neg(A \vee B)\big) \vee \neg(A \vee B \mathrel{\Box\!\!\rightarrow} \neg A)$$

but then again one literal from the principal formulae of a rule $\mathsf{R}_{m,n}$ would be replaced by *two* literals with $\Box\!\!\rightarrow$ as the main connective and we would not be able to derive a single formula with $\Box\!\!\rightarrow$ as the main connective. It is not clear whether the method can somehow be adapted to work in this case as well.

So let us consider the strong counterfactual implication. Using the methods of Chapter 3 on the translation axiom given in (5.1) yields the two rules R_{t1} and R_{t2} given by

$$\begin{aligned}\mathsf{R}_{t1} :=& \mathcal{P}_t/(p \preccurlyeq q), (r \mathrel{\Box\!\!\Rightarrow} s) \Rightarrow\\ \mathsf{R}_{t2} :=& \mathcal{P}_t/\ \Rightarrow (p \preccurlyeq q), (r \mathrel{\Box\!\!\Rightarrow} s)\end{aligned}$$

for the premisses

$$\mathcal{P}_t = \left\{ \begin{array}{l} (p \Rightarrow r; \mathcal{C}_\emptyset), (p, s \Rightarrow\ \mathcal{C}_\emptyset), (r \Rightarrow p, s; \mathcal{C}_\emptyset),\\ (q \Rightarrow r; \mathcal{C}_\emptyset), (q \Rightarrow s; \mathcal{C}_\emptyset), (r, s \Rightarrow q; \mathcal{C}_\emptyset) \end{array} \right\}$$

Moreover, by construction we have equivalence of the translation rules with the translation axioms over $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$. This gives us soundness and completeness of the rule sets from Proposition 5.3.1 not involving absoluteness extended with the translation rules.

Proposition 5.4.2 (c)**.** *The following calculi are sound and complete for the specified logics:*

$\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R}_{t1}\mathsf{R}_{t2}\mathsf{CutCon}$	*for* $\mathbb{V}_{\preccurlyeq,\Box\Rightarrow}$		
$\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R}_{\mathsf{N}}\mathsf{R}_{t1}\mathsf{R}_{t2}\mathsf{CutCon}$	*for* $\mathbb{VN}_{\preccurlyeq,\Box\Rightarrow}$	$\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R}_{\mathsf{T}}\mathsf{R}_{t1}\mathsf{R}_{t2}\mathsf{CutCon}$	*for* $\mathbb{VT}_{\preccurlyeq,\Box\Rightarrow}$
$\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R}_{\mathsf{T}}\mathsf{R}_{\mathsf{W}}\mathsf{R}_{t1}\mathsf{R}_{t2}\mathsf{CutCon}$	*for* $\mathbb{VW}_{\preccurlyeq,\Box\Rightarrow}$	$\mathsf{Gc}\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathsf{R}_{\mathsf{C}}\mathsf{R}_{t1}\mathsf{R}_{t2}\mathsf{CutCon}$	*for* $\mathbb{VC}_{\preccurlyeq,\Box\Rightarrow}$

Proof. Applying the translation procedure from Theorem 3.2.14 to the axioms $(p \mathrel{\Box\!\!\Rightarrow} q) \to \neg((p \wedge \neg q) \preccurlyeq (p \wedge q))$ and $\neg((p \wedge \neg q) \preccurlyeq (p \wedge q)) \to (p \mathrel{\Box\!\!\Rightarrow} q)$ yields exactly the rules R_{t1} and R_{t2}. Thus by the same theorem each of the rules is equivalent to the original axiom over $\mathcal{R}_{\mathbb{V}_\preccurlyeq}$. This together with soundness and completeness of the rule sets in the language with $\preccurlyeq$ (Proposition 5.3.1) and the fact that the connective $\Box\!\!\Rightarrow$ is defined by the translation axioms gives the stated result. □

Again we now might exploit the representation of rules in the cut closures of these rule sets to show decidability of the logics. For this we need to show that the cut closures of the rule sets given above are contraction closed. We do this by considering the representation of such rules in terms of cut graphs.

Lemma 5.4.3 (c)**.** *Let $\mathcal{L}$ be one of $\mathbb{V}_{\Box\Rightarrow}, \mathbb{VN}_{\Box\Rightarrow}, \mathbb{VT}_{\Box\Rightarrow}, \mathbb{VW}_{\Box\Rightarrow}$ or $\mathbb{VC}_{\Box\Rightarrow}$ and let $\mathcal{R}$ be the corresponding set of modal rules as given in Proposition 5.4.2. Then the rule set* $\mathsf{cc}\left(\mathcal{R}_{\mathbb{V}_\preccurlyeq}\mathcal{R}\mathsf{R}_{t1}\mathsf{R}_{t2}\right)$

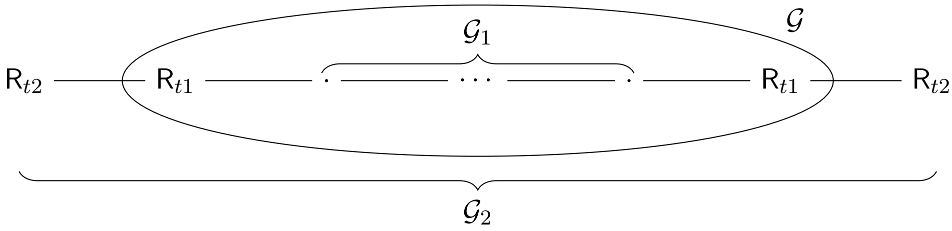


Figure 5.5: The cut graphs constructed in the proof of Lemma 5.4.3

is contraction closed.

Proof. In view of Theorem 4.1.29 we only need to check that contractions of path rules in $\mathsf{cc}\left(\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathcal{R}\mathsf{R}_{t1}\mathsf{R}_{t2}\right)$ are subsumed. So let R be such a path rule. Then the cut graph $\mathcal{G}$ for its representing cut tree is a path. On the other hand we know that the only rule in $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathcal{R}\mathsf{R}_{t1}\mathsf{R}_{t2}$ in which the connective $\Box\!\!\Rightarrow$ occurs on the left hand side of the principal formulae is the rule R_{t1} and likewise the only rule in which it occurs on the right hand side is R_{t2}. Thus whenever one of these rules, say R_{t1}, appears in the cut graph for R but is not one of the two endpoints, then one of the adjacent nodes must be the other of the two rules, in this case R_{t2}, and a formula $(p \Box\!\!\Rightarrow q)$ must appear in the principal formulae of both rules. But as is easy to check the rule $\mathsf{cut}(\mathsf{R}_{t2}, \mathsf{R}_{t1}, (p \Box\!\!\Rightarrow q))$ is subsumed by the congruence rule and thus also by the rule $\mathsf{R}_{1,1}$. This means that w.l.o.g. we may assume that if the rules R_{t1} or R_{t2} occur in the cut graph $\mathcal{G}$, then they occur as one of its two endpoints. Now suppose we have a contraction of R on literals occurring in the principal formulae of the rules at the two endpoints of $\mathcal{G}$. If the literals are of the form $(r \preccurlyeq s)$, then the connective $\Box\!\!\Rightarrow$ does not occur at all in $\mathcal{G}$, and we are done using Proposition 5.3.1. If on the other hand the contracted formula is $(r \Box\!\!\Rightarrow s)$ for some variables r and s, then the endpoints of $\mathcal{G}$ must be both R_{t1} or both R_{t2}. We consider the case that they are both R_{t1}, the other case is analogous. Let $\mathcal{G}_1$ be the cut graph constructed from $\mathcal{G}$ by deleting the two endpoints, and let $\mathcal{G}_2$ be the one constructed from $\mathcal{G}$ by appending nodes R_{t2} to each of the endpoints of $\mathcal{G}$ (see Figure 5.5). Furthermore, let R_1 and R_2 be rules such that $\mathcal{G}_1$ is the cut graph for the cut tree representing R_1 and $\mathcal{G}_2$ is the cut graph for the cut tree representing R_2. Then again since the rule $\mathsf{cut}(\mathsf{R}_{t2}, \mathsf{R}_{t1}, (p \Box\!\!\Rightarrow q))$ is subsumed by $\mathsf{R}_{1,1}$ the rule R_1 subsumes R_2 and vice versa. Now if a contraction of the rule R_1 on the literals not occurring in the rule R is subsumed by a rule in our rule set, then obviously the original contraction of R is subsumed. Furthermore, if the latter is the case, then the corresponding contraction of rule R_2 is subsumed. Since R_2 subsumes R_1 this means that the contraction of R_1 is subsumed by a rule in the rule set if and only if the original contraction of R is subsumed by a rule in the rule set. But by the reasoning above the connective $\Box\!\!\Rightarrow$ does not occur in $\mathcal{G}_1$, and thus using Proposition 5.3.1 we know that contractions of this rule are subsumed by rules in our rule set. Therefore the original contraction of R is subsumed by a rule in our rule set as well. $\square$

$$\Pi_n \Rightarrow \Sigma_m := (r_1 \mathrel{\Box\!\!\Rightarrow} s_1), \dots, (r_n \mathrel{\Box\!\!\Rightarrow} s_n) \Rightarrow (p_1 \mathrel{\Box\!\!\Rightarrow} q_1), \dots, (p_m \mathrel{\Box\!\!\Rightarrow} q_m)$$

$$\mathsf{R}'_{n,m} := \mathcal{P}'_{n,m} / \Pi_n \Rightarrow \Sigma_m \text{ where}$$

$$\mathcal{P}'_{n,m} := \begin{array}{l} \big\{(r_k, s_k, \vec{s}_I \Rightarrow \vec{r}_{[n]\smallsetminus I}, \vec{p}_J, \vec{q}_{[m]\smallsetminus J}; \mathcal{C}_\emptyset) \mid k \leq n,\ I \subseteq [n],\ J \subseteq [m]\big\} \\ \cup \big\{(p_k, \vec{s}_I \Rightarrow q_k, \vec{r}_{[n]\smallsetminus I}, \vec{p}_J, \vec{q}_{[k-1]\smallsetminus J}; \mathcal{C}_\emptyset) \mid I \subseteq [n], k \leq m, J \subseteq [k-1]\big\} \end{array}$$

$$\mathsf{T}'_m := \mathcal{P}'^{\mathsf{T}}_m / \Pi_0 \Rightarrow \Sigma_m \text{ where}$$

$$\mathcal{P}'^{\mathsf{T}}_m := \begin{array}{l} \big\{(\Rightarrow \vec{p}_J, \vec{q}_{[m]\smallsetminus J}; \mathcal{C}_{\mathsf{id}}) \mid J \subseteq [m]\big\} \\ \cup \big\{(p_k, \Rightarrow q_k, \vec{p}_J, \vec{q}_{[k-1]\smallsetminus J}; \mathcal{C}_\emptyset) \mid k \leq m, J \subseteq [k-1]\big\} \end{array}$$

$$\mathsf{W}'_{n,m} := \mathcal{P}'^{\mathsf{W}}_{n,m} / \Pi_n \Rightarrow \Sigma_m \text{ where}$$

$$\mathcal{P}'^{\mathsf{W}}_{n,m} := \begin{array}{l} \big\{(\vec{s}_I \Rightarrow \vec{r}_{[n]\smallsetminus I}, \vec{p}_J, \vec{q}_{[m]\smallsetminus J}; \mathcal{C}_{\mathsf{id}} \mid I \subseteq [n], J \subseteq [m]\big\} \\ \cup \big\{(p_k, \vec{s}_I \Rightarrow q_k, \vec{r}_{[n]\smallsetminus I}, \vec{p}_J, \vec{q}_{[k-1]\smallsetminus J}; \mathcal{C}_\emptyset) \mid I \subseteq [m], k \leq n, J \subseteq [k-1]\big\} \end{array}$$

$$\mathsf{R}'_{\mathsf{W2}} := \{(\Rightarrow p; \mathcal{C}_{\mathsf{id}}), (q \Rightarrow ; \mathcal{C}_{\mathsf{id}})\} / (p \mathrel{\Box\!\!\Rightarrow} q) \Rightarrow$$
$$\mathsf{R}'_{\mathsf{C2}} := \{(\Rightarrow p; \mathcal{C}_{\mathsf{id}}), (p \Rightarrow q; \mathcal{C}_{\mathsf{id}})\} / \ \Rightarrow (p \mathrel{\Box\!\!\Rightarrow} q)$$

$$\mathcal{R}_{\mathbb{V}_{\Box\!\Rightarrow}} := \big\{\mathsf{R}'_{n,m} \mid n \geq 0, m \geq 1\big\}$$
$$\mathcal{R}_{\mathbb{VN}_{\Box\!\Rightarrow}} := \big\{\mathsf{R}'_{n,m} \mid n + m \geq 1\big\} \qquad \mathcal{R}_{\mathbb{VW}_{\Box\!\Rightarrow}} := \big\{\mathsf{W}'_{n,m} \mid n + m \geq 1\big\}$$
$$\mathcal{R}_{\mathbb{VT}_{\Box\!\Rightarrow}} := \mathcal{R}_{\mathbb{V}_{\Box\!\Rightarrow}} \cup \{\mathsf{T}'_m \mid m \geq 1\} \qquad \mathcal{R}_{\mathbb{VC}_{\Box\!\Rightarrow}} := \mathcal{R}_{\mathbb{V}_{\Box\!\Rightarrow}} \cup \{\mathsf{R}'_{W2}, \mathsf{R}'_{C2}\}$$

Table 5.5: The rules and rule sets for strong conditional logics formulated using $\mathrel{\Box\!\!\Rightarrow}$.

While this yields decidability results analogous to the ones in Theorem 5.3.2, we can also use this result to explicitly construct sequent calculi in the language with $\mathrel{\Box\!\!\Rightarrow}$ as the only connective. For the basic logic $\mathbb{V}$ this is done by explicitly computing the cut between a rule $\mathsf{R}_{m,n}$ and m instances of the rule R_{t1} as well as n instances of the rule R_{t2}. Similarly, if $\mathcal{L}$ is one of the other logics we compute the cuts between rules in $\mathcal{R}_{\mathcal{L}_{\preccurlyeq}}$ and the appropriate number of instances of R_{t1} and R_{t2}. This gives the rules described in the following definition.

Definition 5.4.4. For variables $p_1, \dots, p_m$ and a set $I \subseteq \{1, \dots, n\}$ of indices we write $\vec{p}_I$ for the sequent consisting of all variables p_i with $i \in I$, and for $n \in \mathbb{N}$ we abbreviate $\{0, \dots, n\}$ to $[n]$. Then the rules $\mathsf{R}'_{n,m}, \mathsf{T}'_m, \mathsf{W}'_{n,m}, \mathsf{R}'_{W2}$ and R'_{C2} as well as the rule sets $\mathcal{R}_{\mathcal{L}}$ for $\mathcal{L} \in \{\mathbb{V}_{\Box\!\Rightarrow}, \mathbb{VN}_{\Box\!\Rightarrow}, \mathbb{VT}_{\Box\!\Rightarrow}, \mathbb{VW}_{\Box\!\Rightarrow}, \mathbb{VC}_{\Box\!\Rightarrow}\}$ are defined as shown in Table 5.5.

Then while the rules are by construction sound, we can prove completeness similarly to the proof of Theorem 5.3.6 by showing that all the rules in the cut closure of the respective rule set are subsumed by the constructed rules. Since the calculus given by the cut closure has cut elimination, this gives cut-free completeness of the calculus given by the constructed rules.

Theorem 5.4.5 (c)**.** *Let $\mathcal{L}$ be a logic in $\{\mathbb{V}_{\Box\!\Rightarrow}, \mathbb{VN}_{\Box\!\Rightarrow}, \mathbb{VT}_{\Box\!\Rightarrow}, \mathbb{VW}_{\Box\!\Rightarrow}, \mathbb{VC}_{\Box\!\Rightarrow}\}$. Then the calculus* $\mathsf{Gc}\mathcal{R}_{\mathcal{L}}\mathsf{Con}$ *is sound and complete for $\mathcal{L}$.*

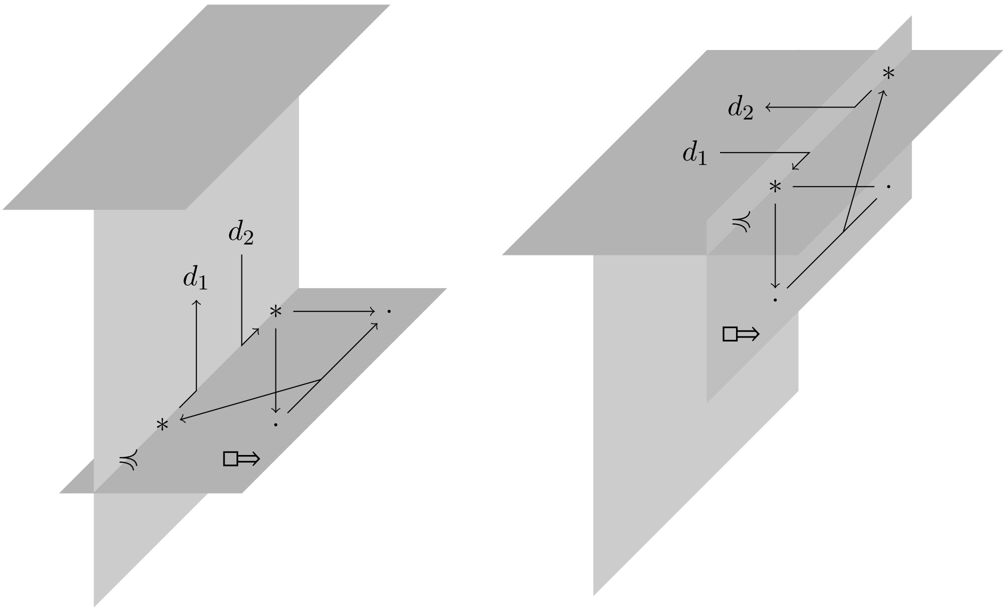


Figure 5.6: Cutting with the translation rules in the proof of Theorem 5.4.5

Proof. For soundness we first check that indeed the rules of $\mathcal{R}_{\mathbb{V}_{\Box\!\Rightarrow}}$ arise by cutting on every principal formula of a rule in $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$ with the translation rules. Such a situation is shown schematically in Figure 5.6. Here some of the premisses of the translation rules have been omitted for clarity. This is no problem, since the omitted premisses do not contribute to the premisses of the resulting rule. It can be seen that when cutting on a negative literal $(p_k \preccurlyeq q_k)$ this literal is replaced by a positive literal $(p'_k \Box\!\!\Rightarrow q'_k)$ and in every premiss d_1 of $\mathsf{R}_{m,n}$ emerging from the variable p_k the latter is replaced by a tail emerging from p'_k and a head pointing to q'_k. On the other hand, every premiss with a head pointing to q_k is replaced by two premisses pointing to p'_k resp. q'_k instead. Similarly for cuts on positive literals. Thus cutting on all literals of $\mathsf{R}_{m,n}$ yields the rule $\mathsf{R}'_{n,m}$. Since the translation rules as well as the rules $\mathsf{R}_{m,n}$ are sound for $\mathbb{V}_{\preccurlyeq,\Box\!\Rightarrow}$ this gives soundness of $\mathsf{R}'_{n,m}$ using Lemma 2.4.5. The reasoning for the other rules is similar, where in the case of the rule $\mathsf{R}'_{\mathsf{C2}}$ in the presence of the premisses $(\Rightarrow p; \mathcal{C}_{\mathsf{id}})$ and $(p \Rightarrow q; \mathcal{C}_{\mathsf{id}})$ we may safely omit the premiss $(\Rightarrow q; \mathcal{C}_{\mathsf{id}})$ since it is derivable from the other two premisses using cut.

For completeness we make use of the fact that if for one of the logics under consideration the set $\mathcal{R}$ is the corresponding set of modal rules as given in Proposition 5.4.2, and if a sequent $\Gamma \Rightarrow \Delta \in \mathcal{S}(\mathcal{F}(\Box\!\!\Rightarrow))$ is cut-free derivable in $\mathsf{Gc}\,\mathsf{cc}\left(\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathcal{R}\mathsf{R}_{t1}\mathsf{R}_{t2}\right)\mathsf{Con}$, then it has a derivation in which the connective $\preccurlyeq$ does not occur. This holds since rules with restrictions and therefore also cuts between rules with restrictions have the subformula property, and thus if the connective $\preccurlyeq$ would occur in such a derivation, then it would also occur in the sequent $\Gamma \Rightarrow \Delta$. Thus it suffices to show that the rules in $\mathsf{cc}\left(\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathcal{R}\mathsf{R}_{t1}\mathsf{R}_{t2}\right)$ whose principal formulae do not mention the connective $\preccurlyeq$ are derivable rules in $\mathcal{R}_{\mathcal{L}_{\Box\!\Rightarrow}}$. This is straightforward for the rules with only one formula in the principal formulae. So let R be a rule in $\mathsf{cc}\left(\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathcal{R}\mathsf{R}_{t1}\mathsf{R}_{t2}\right)$ with at least two formulae in the principal formulae and let $\mathcal{G}_R$ be the cut graph for a cut tree representing R. Then as in the proof of Lemma 5.4.3 if in $\mathcal{G}_R$ we have a cut between the rules R_{t1} and R_{t2} on a literal $(p \mathrel{\Box\!\!\Rightarrow} q)$ we may replace this cut by the rule $\mathsf{R}_{1,1}$. Thus w.l.o.g. we may assume that all cuts in the construction of R are on formulae with main connective $\preccurlyeq$. This means that nodes in $\mathcal{G}_R$ which are labelled with a rule R_{t1} or R_{t2} must have degree one. Omitting all these nodes yields a cut graph $\mathcal{G}_Q$ for the cut tree representing a rule Q from $\mathsf{cc}\left(\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathcal{R}\right)$. But we know from the proof of Theorem 5.3.6 that such a rule is subsumed by a rule Q' in the rule set $\mathcal{R}_{\mathcal{L}_{\preccurlyeq}}$ in case $\mathcal{L} \neq \mathbb{VC}$ and a derivable rule in case $\mathcal{L} = \mathbb{VC}$. Thus for $\mathcal{L} \neq \mathbb{VC}$ we simply replace the subgraph $\mathcal{G}_Q$ in $\mathcal{G}_R$ by the cut graph $\mathcal{G}_{Q'}$ for a cut tree representing the rule Q'. This yields a cut graph $\mathcal{G}_{R'}$ for a rule constructed from a rule in $\mathcal{R}_{\mathcal{L}_{\preccurlyeq}}$ by cutting on all principal formulae with the translation rules. Since this is exactly how the rules in $\mathcal{R}'_{\mathcal{L}_{\Box\Rightarrow}}$ are constructed, this rule is subsumed by a rule in $\mathcal{R}'_{\mathcal{L}_{\Box\Rightarrow}}$. In case $\mathcal{L} = \mathbb{VC}$ similar as in the case for $\preccurlyeq$ if the cut graph $\mathcal{G}_Q$ involves a cut between a rule $\mathsf{R}_{m,n}$ and R_{C} the resulting rule is derivable using a number of applications of $\mathsf{R}'_{\mathsf{W2}}$ followed by a number of applications of $\mathsf{R}'_{\mathsf{C2}}$. Thus in any case the original rule R is a derivable rule in $\mathcal{R}'_{\mathcal{L}_{\Box\Rightarrow}}$, and thus we obtain completeness of $\mathsf{Gc}\mathcal{R}'_{\mathcal{L}_{\Box\Rightarrow}}\mathsf{Con}$ for $\mathcal{L}_{\Box\Rightarrow}$. $\square$

Again this immediately yields complexity results for the logics under scrutiny.

Corollary 5.4.6 (c)**.** *For $\mathcal{L} \in \{\mathbb{V}_{\Box\Rightarrow}, \mathbb{VN}_{\Box\Rightarrow}, \mathbb{VT}_{\Box\Rightarrow}, \mathbb{VW}_{\Box\Rightarrow}, \mathbb{VC}_{\Box\Rightarrow}\}$ the validity problem for $\mathcal{L}$ is in* PSPACE.

Proof. Let $\mathcal{L}$ be such a logic. Inspection of the rules shows that the rule set $\mathcal{R}'_{\mathcal{L}}$ is contraction closed and tractable. Thus using modified applications of the rules (Definition 2.6.1) and the procedure from the proof of Theorem 2.7.8 we obtain the complexity bounds. $\square$

5.5 Interpolation

Given a cut-free sequent system which is sound and complete for a logic, we often can use this system to establish interpolation results for this logic. This is the case for the calculi for the conditional logics considered as well. We begin by recalling the basic notions.

Definition 5.5.1 (c)**.** Let $\mathcal{L}$ be a logic based on classical propositional logic. Given two formulae A, B a formula I is called an *$\mathcal{L}$-interpolant for A and B* if

1. $\models_{\mathcal{L}} A \to I$ and $\models_{\mathcal{L}} I \to B$
2. I satisfies the *variable condition*, i.e., $\mathsf{var}(I) \subseteq \mathsf{var}(A) \cap \mathsf{var}(B)$.

We say that the logic $\mathcal{L}$ has the *Craig interpolation property* (or *interpolation property* for short) if whenever for two formulae A, B we have $\models_{\mathcal{L}} A \to B$, then there exists an $\mathcal{L}$-interpolant for A and B.

One standard method of obtaining the interpolation property for a logic is by lifting the notion of an interpolant to the level of sequents using so-called split sequents, and employ an induction on the depth of the derivation in a cut-free sequent calculus, see e.g. [TS00] for the propositional and first-order case.

Definition 5.5.2 (c)**.** A *split sequent* is a tuple of sequents $\Gamma_1 \Rightarrow \Delta_1$ and $\Gamma_2 \Rightarrow \Delta_2$ which we write as $\Gamma_1 \mid \Gamma_2 \Rightarrow \Delta_1 \mid \Delta_2$. We then say that $\Gamma_1 \mid \Gamma_2 \Rightarrow \Delta_1 \mid \Delta_2$ is a *splitting* of $\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2$. If $\mathcal{R}$ is a set of rules with restrictions, we say that a formula I is an *$\mathcal{R}$-interpolant* for the split sequent $\Gamma_1 \mid \Gamma_2 \Rightarrow \Delta_1 \mid \Delta_2$ if

1. $\vdash_{\mathcal{R}\mathsf{Con}} \Gamma_1 \Rightarrow \Delta_1, I$ and $\vdash_{\mathcal{R}\mathsf{Con}} I, \Gamma_2 \Rightarrow \Delta_2$
2. $\mathsf{var}\,(I) \subseteq \mathsf{var}\,(\Gamma_1 \Rightarrow \Delta_1) \cap \mathsf{var}\,(\Gamma_2 \Rightarrow \Delta_2)$.

A sequent *admits interpolation in* $\mathcal{R}$ if all its splittings have an interpolant in $\mathcal{R}$. An application of a rule R from $\mathcal{R}$ *supports interpolation in* $\mathcal{R}$ if whenever all its premisses admit interpolation in $\mathcal{R}$, then so does its conclusion.

Note that we allow Contraction in clause 1 of the above definition. If all rules of a sequent calculus admit interpolation it is possible to show that all derivable sequents admit interpolation as well. Provided the calculus is sound and complete for a logic this then yields the interpolation property for that logic.

Lemma 5.5.3 (c)**.** *Let $\mathcal{R}$ be a set of rules with restrictions. If all the rules in $\mathcal{R}$ support interpolation in $\mathcal{R}$, then every sequent derivable in $\mathcal{R}\mathsf{Con}$ admits interpolation. Thus if furthermore $\to_{\mathsf{R}}$ is an invertible rule in $\mathcal{R}\mathsf{Con}$ and $\mathcal{R}\mathsf{Con}$ is sound and complete for a logic $\mathcal{L}$, then $\mathcal{L}$ has the Craig interpolation property.*

Proof. The first claim follows by a straightforward induction on the depth of the derivation in $\mathcal{R}$. The base cases are analogous to the propositional case [TS00] and the induction step follows since Con and every rule in $\mathcal{R}$ support interpolation. The second claim follows since if $\mathcal{R}\mathsf{Con}$ is sound and complete for the logic $\mathcal{L}$ and we have $\models_{\mathcal{L}} A \to B$ we obtain $\vdash_{\mathcal{R}\mathsf{Con}} \Rightarrow A \to B$ and thus with invertibility of the rule $\to_{\mathsf{R}}$ also $\vdash_{\mathcal{R}\mathsf{Con}} A \Rightarrow B$. Since by the first claim this sequent admits interpolation there is a $\mathcal{R}$-interpolant for the split sequent $A \mid \emptyset \Rightarrow \emptyset \mid B$. By soundness of $\mathcal{R}\mathsf{Con}$ this is also an $\mathcal{L}$-interpolant for A and B. $\square$

We now use this method to show that almost all of the conditional logics considered in this chapter have the interpolation property. The main difficulty lies in showing that the rules in $\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$ support interpolation.

Theorem 5.5.4 (c)**.** *The logic $\mathbb{V}_{\preccurlyeq}$ has the Craig interpolation property.*

Proof. We need to show that all rules in $\mathsf{Gc}\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}$ support interpolation. For the propositional rules this is standard [TS00]. For an application of a conditional rule $\mathsf{R}_{m,n}$ we need to construct an interpolant for a splitting of the conclusion from the interpolants of the corresponding splittings of the premisses. We first deal with the case $n = 2$. So suppose the application of $\mathsf{R}_{m,2}$ has the conclusion

$$(A_1 \preccurlyeq B_1), \ldots, (A_m \preccurlyeq B_m) \Rightarrow (C_1 \preccurlyeq D_1), (C_2 \preccurlyeq D_2)$$

and premisses

$$\begin{gathered}\{D_1 \Rightarrow B_1, \ldots, B_m, C_1, C_2\} \cup \{D_2 \Rightarrow B_1, \ldots, B_m, C_1, C_2\} \\ \cup \{A_k \Rightarrow B_1, \ldots, B_{k-1}, C_1, C_2 \mid 1 \leq k \leq m\}\end{gathered}$$

and furthermore suppose that we have a splitting of the conclusion. In a first step we assume that the splitting separates the two positive formulae and alternates on the negative formulae, i.e., that it has the form $\Gamma_1 \mid \Gamma_2 \Rightarrow \Delta_1 \mid \Delta_2$ with

$$\begin{aligned}\Gamma_1 &= \{(A_i \preccurlyeq B_i) \mid 1 \leq i \leq m,\ i \text{ odd}\} & \Gamma_2 &= \{(A_i \preccurlyeq B_i) \mid 1 \leq i \leq m,\ i \text{ even}\} \\ \Delta_1 &= (C_1 \preccurlyeq D_1) & \Delta_2 &= (C_2 \preccurlyeq D_2)\end{aligned} \tag{5.2}$$

Then for k with $1 \leq k \leq m$ let I_k be the interpolant for the corresponding splitting of the premiss $A_k \Rightarrow B_1, \ldots, B_{k-1}, C_1, C_2$, i.e., for the split sequent

$$A_k \mid \emptyset \Rightarrow \{B_i \mid 1 \leq i \leq k-1,\ i \text{ odd}\}, C_1 \mid \{B_i \mid 1 \leq i \leq k-1,\ i \text{ even}\}, C_2$$

if k is odd, and for the split sequent

$$\emptyset \mid A_k \Rightarrow \{B_i \mid 1 \leq i \leq k-1,\ i \text{ odd}\}, C_1 \mid \{B_i \mid 1 \leq i \leq k-1,\ i \text{ even}\}, C_2$$

if k is even. Similarly, for $k \in \{1, 2\}$ let J_k be the interpolant for the corresponding splitting of the premiss $D_k \Rightarrow B_1, \ldots, B_m, C_1, C_2$. Now for every odd k with $1 \leq k \leq m$ we define the formulae X_k, Y_k, Z_k and V_k, W_k by

$$X_k := \bigvee\nolimits_{1 \leq \ell \leq k,\ \ell \text{ odd}} I_\ell \qquad Z_k := J_1 \vee \bigvee\nolimits_{k < \ell \leq m,\ \ell \text{ odd}} I_\ell$$

$$Y_k := \begin{cases} \neg I_{k+1} \vee \neg J_2 & \text{if } k = \max\{\ell \mid 1 \leq \ell \leq m,\ \ell \text{ odd}\} \text{ and } k \neq m \\ \neg J_2 & \text{if } k = m \\ \neg I_{k+1} & \text{otherwise} \end{cases}$$

$$V_k := (X_k \preccurlyeq Y_k) \qquad W_k := (Y_k \preccurlyeq Z_k)$$

Then we can show the following.

Claim 1: For every odd k with $1 \leq k \leq m$ we have $\vdash_{\mathsf{Gc}\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathsf{Con}} \Gamma_1, W_k \Rightarrow \Delta_1, V_k$.

This can be seen by inserting W_k instead of the literal $(A_{k+1} \preccurlyeq B_{k+1})$ and V_k instead of the literal $(C_2 \preccurlyeq D_2)$ into the rule pattern, checking that all necessary premisses are derivable and applying the rule. Thus e.g. in case $k < m-1$ we derive the necessary premisses as follows. For the premiss emerging from D_1 we have

$$\dfrac{\dfrac{D_1 \Rightarrow B_1, B_3, \ldots, B_k, J_1, B_{k+2}, \ldots, B_m, C_1}{D_1 \Rightarrow B_1, B_3, \ldots, B_k, J_1 \vee \bigvee_{k<\ell\leq m,\ \ell \text{ odd}} I_\ell, B_{k+2}, \ldots, B_m, C_1}\ \mathsf{W}, \vee_{\mathsf{R}}}{D_1 \Rightarrow B_1, B_3, \ldots, B_k, \underbrace{J_1 \vee \bigvee_{k<\ell\leq m,\ \ell \text{ odd}} I_\ell}_{Z_k}, B_{k+2}, \ldots, B_m, C_1, X_k}\ \mathsf{W}$$

The premiss emerging from Y_k is derived using

$$\dfrac{\dfrac{\Rightarrow I_{k+1}, B_1, B_3, \ldots, B_k, C_1}{\neg I_{k+1} \Rightarrow B_1, B_3, \ldots, B_k, C_1}\ \neg\mathsf{L}}{\underbrace{\neg I_{k+1}}_{Y_k} \Rightarrow B_1, B_3, \ldots, B_k, Z_k, B_{k+1}, \ldots, B_m, C_1, Y_k}\ \mathsf{W}$$

For odd i with $i \leq k$ the premiss emerging from A_i is derived as in

$$\dfrac{A_k \Rightarrow B_1, B_3, \ldots, B_{k-2}, C_1, I_k}{A_k \Rightarrow B_1, B_3, \ldots, B_{k-2}, C_1, \underbrace{\bigvee_{\ell\leq k} I_\ell}_{X_k}}\ \mathsf{W}, \vee_{\mathsf{R}}$$

In order to derive the premiss emerging from the lower occurrence of Y_k we use:

$$\dfrac{\dfrac{\Rightarrow I_{k+1}, B_1, B_3, \ldots, B_k, C_1}{\neg I_{k+1} \Rightarrow B_1, B_3, \ldots, B_k, C_1}\ \neg\mathsf{L}}{\underbrace{\neg I_{k+1}}_{Y_k} \Rightarrow B_1, \mathsf{B}_3, \ldots, B_k, C_1, X_k}\ \mathsf{W}$$

And finally for odd i with $k < i \leq m$ we derive the premiss emerging from A_i as in

$$\dfrac{\dfrac{A_{k+2} \Rightarrow I_{k+2}, B_1, B_3, \ldots, B_k, C_1}{A_{k+2} \Rightarrow J_1 \vee \bigvee_{k<\ell\leq m,\ \ell \text{ odd}} I_\ell, B_1, B_3, \ldots, B_k, C_1}\ \mathsf{W}, \vee_{\mathsf{R}}}{A_{k+2} \Rightarrow \underbrace{J_1 \vee \bigvee_{k<\ell\leq m,\ \ell \text{ odd}} I_\ell}_{Z_k}, B_1, B_3, \ldots, B_k, C_1, X_k}\ \mathsf{W}$$

Thus we may apply rule $\mathsf{R}_{\lceil \frac{m}{2} \rceil + 1, 2}$ to derive $\Gamma_1, (Y_k \preccurlyeq Z_k) \Rightarrow \Delta_1, (X_k \preccurlyeq Y_k)$. The cases that $k \in \{m-1, m\}$ are similar. This proves Claim 1.

Claim 2: For every partition (F, S) of $\{k \mid 1 \leq k \leq m,\ k \text{ odd}\}$ we have $\vdash_{\mathsf{Gc}\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathsf{Con}} \Gamma_2, \{V_k \mid k \in F\} \Rightarrow \Delta_2, \{W_k \mid k \in S\}$.

This is shown similar to above by inserting for $k \in F$ the V_k instead of the $(A_k \preccurlyeq B_k)$

and for $k \in S$ the W_k as positive literals in place of $(C_1 \preccurlyeq D_1)$ into the rule pattern. Then again we check that all necessary premisses are derivable and apply the rule $\mathsf{R}_{|F|+\lfloor \frac{m}{2} \rfloor, |S|+1}$. So suppose we have an odd k with $k \in F$. Then we derive the premiss emerging from X_k by

$$\dfrac{\dfrac{\dfrac{I_1 \Rightarrow C_2}{I_1 \Rightarrow \{B_i \mid 1 \leq i \leq k-1,\ i \text{ even}\}, C_2}\ \mathsf{W} \qquad \dots \qquad I_k \Rightarrow \{B_i \mid 1 \leq i \leq k-1,\ i \text{ even}\}, C_2}{\bigvee_{1 \leq \ell \leq k,\ \ell \text{ odd}} I_\ell \Rightarrow \{B_i \mid 1 \leq i \leq k-1,\ i \text{ even}\}, C_2}\ \vee_{\mathsf{L}}}{\underbrace{\bigvee_{1 \leq \ell \leq k,\ \ell \text{ odd}} I_\ell}_{X_k} \Rightarrow \{B_i \mid 1 \leq i \leq k-1,\ i \text{ even}\}, C_2, \{Y_\ell \mid \ell \in S\}}\ \mathsf{W}$$

Moreover, for every even $k < m$ we derive the premiss emerging from A_k by

$$\dfrac{A_k \Rightarrow B_2, B_4, \dots, B_{k-2}, \overbrace{\neg I_k}^{Y_{k-1}}, C_2}{A_k \Rightarrow B_2, B_4, \dots, B_{k-2}, \{Y_i \mid i \leq m,\ i \text{ odd}\}, C_2}\ \mathsf{W}$$

and in case m is even we have

$$\dfrac{A_m \Rightarrow B_2, B_4, \dots, B_{m-2}, \neg I_m, C_2}{A_m \Rightarrow B_2, \mathsf{B}_4, \dots, B_{m-2}, \{Y_i \mid i < m-1,\ i \text{ odd}\}, \underbrace{\neg I_m \vee \neg J_2}_{Y_{m-1}}, C_2}\ \mathsf{W}, \vee_{\mathsf{R}}$$

For the premiss emerging from D_2 in case $\max\{i \mid i \leq m,\ i \text{ odd}\} \neq m$ we have

$$\dfrac{\dfrac{D_2 \Rightarrow \{B_i \mid 1 \leq i \leq m,\ i \text{ even}\}, C_2, \neg J_2}{D_2 \Rightarrow \{B_i \mid 1 \leq i \leq m,\ i \text{ even}\}, C_2, \neg I_m \vee \neg J_2}\ \mathsf{W}, \vee_{\mathsf{R}}}{D_2 \Rightarrow \{B_i \mid 1 \leq i \leq m,\ i \text{ even}\}, C_2, \underbrace{\neg I_m \vee \neg J_2}_{Y_{\max\{i \mid i \leq m,\ i \text{ odd}\}}}, \{Y_i \mid i \leq m-2,\ i \text{ odd}\}}\ \mathsf{W}$$

The case $\max\{i \mid i \leq m,\ i \text{ odd}\} = m$ is similar. Finally, we have for every odd $k \in S$:

$$\dfrac{\dfrac{J_1 \Rightarrow \{B_i \mid 1 \leq i \leq m,\ i \text{ even}\}, C_2 \qquad \dots \qquad I_3 \Rightarrow B_2, C_2 \qquad I_1 \Rightarrow C_2}{J_1 \vee \bigvee_{k < \ell \leq m,\ \ell \text{ odd}} I_\ell \Rightarrow \{B_i \mid 1 \leq i \leq m,\ i \text{ even}\}, C_2}\ \mathsf{W}, \vee_{\mathsf{L}}}{\underbrace{J_1 \vee \bigvee_{k < \ell \leq m,\ \ell \text{ odd}} I_\ell}_{Z_k} \Rightarrow \{B_i \mid 1 \leq i \leq m,\ i \text{ even}\}, C_2, \{Y_\ell \mid \ell \in S\}}\ \mathsf{W}$$

Thus we have all the necessary premisses to apply the rule $\mathsf{R}_{|F|+\lfloor \frac{m}{2} \rfloor, |S|+1}$ and obtain the sequent $\Gamma_2, \{V_k \mid k \in F\} \Rightarrow \Delta_2, \{W_k \mid k \in S\}$. This shows Claim 2.

Now we define the interpolant I by

$$I := \bigwedge_{1 \leq k \leq m,\ k \text{ odd}} (\neg W_k \vee V_k)\ .$$

Then using the propositional rules from Claim 1 we get $\vdash_{\mathsf{Gc}\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathsf{Con}} \Gamma_1 \Rightarrow \Delta_1, I$ and from Claim 2 we obtain $\vdash_{\mathsf{Gc}\mathcal{R}_{\mathbb{V}_{\preccurlyeq}}\mathsf{Con}} I, \Gamma_2 \Rightarrow \Delta_2$. Moreover, since the formulae I_k for $k \leq m$ and J_ℓ for $\ell = 1, 2$ were interpolants, their variables occur both in $\Gamma_1 \Rightarrow \Delta_1$ and $\Gamma_2 \Rightarrow \Delta_2$, and since the formula I is constructed from these formulae, it satisfies the variable condition as well. Hence it indeed is an interpolant.

In the next step, assume that (still for an application of $R_{m,2}$ with principal formulae as above) the splitting is $\Gamma_2 \mid \Gamma_1 \Rightarrow \Delta_2 \mid \Delta_1$ with $\Gamma_1, \Gamma_2, \Delta_1, \Delta_2$ as in (5.2) on p. 171 above. Then we simply compute the interpolant I' of the splitting $\Gamma_1 \mid \Gamma_2 \Rightarrow \Delta_1 \mid \Delta_2$ as before and set $I := \neg I'$.

In the case that $n = 1$ this procedure only needs to be slightly adapted. Suppose we again have a splitting $\Gamma_1 \mid \Gamma_2 \Rightarrow \emptyset \mid \Delta_2$ of the principal formulae with Γ_1 and Γ_2 as in (5.2). Then there is no interpolant J_1, since the corresponding premiss does not exists. Instead we simply use the formula $\bot$. In case the splitting is $\Gamma_2 \mid \Gamma_1 \Rightarrow \emptyset \mid \Delta_1$ the premiss $A_1 \Rightarrow C_1$ does not mention any formulae in Γ_2, so omitting I_1 and exchanging the roles of odd and even numbers in the construction of the interpolant as above gives the correct formula. In case we have a splitting $\Gamma_1 \mid \Gamma_2 \Rightarrow \Delta \mid \emptyset$, then again we compute the interpolant I' for the splitting $\Gamma_2 \mid \Gamma_1 \Rightarrow \emptyset \mid \Delta$ and set $I := \neg I'$.

In the most general case in the splitting $\Gamma_1 \mid \Gamma_2 \Rightarrow \Delta_1 \mid \Delta_2$ we have alternating blocks of negative literals, i.e., for some s we have $1 = k_0 \leq k_1 < \cdots < k_s = m$ and

$$\Gamma_1 = \bigcup_{1 \leq i \leq s,\ i \text{ odd}} \Sigma_i \quad \text{and} \quad \Gamma_2 = \bigcup_{1 \leq i \leq s,\ i \text{ even}} \Sigma_i$$

where for $1 \leq i \leq s$ we define

$$\Sigma_i := (A_{k_{i-1}} \preccurlyeq B_{k_{i-1}}), \ldots, (A_{k_i} \preccurlyeq B_{k_i})$$

and two blocks of positive literals, i.e.,

$$\Delta_1 = (C_1 \preccurlyeq D_1), \ldots, (C_\ell \preccurlyeq D_\ell) \quad \text{and} \quad \Delta_2 = (C_{\ell+1} \preccurlyeq D_{\ell+1}), \ldots, (C_n \preccurlyeq D_n)\ .$$

For $1 \leq k \leq m$ we write I'_k for the interpolant of the corresponding splitting of the premiss $A_k \Rightarrow \{B_\ell \mid 1 \leq \ell < k\}, \{C_\ell \mid \ell \leq n\}$ and for $1 \leq k \leq n$ we write J'_k for the one for the premiss $D_k \Rightarrow \{B_\ell \mid 1 \leq \ell \leq m\}, \{C_\ell \mid 1 \leq \ell \leq n\}$. Then in the construction of the interpolant above we replace the formula J_1 by $\bigvee_{(C_\ell \preccurlyeq D_\ell) \in \Delta_1} J'_\ell$ and the formula $\neg J_2$ by $\bigvee_{(C_\ell \preccurlyeq D_\ell) \in \Delta_2} \neg J'_\ell$. The formulae I_ℓ in X_k and Z_k are replaced by $\bigvee_{k_{\ell-1} \leq j \leq k_\ell} I'_j$, and the formulae $\neg I_{\ell+1}$ in Y_k are replaced by $\bigvee_{k_\ell \leq j \leq k_{\ell+1}} \neg I'_j$. Then in the proofs of the claims the formulae W_ℓ and V_ℓ are inserted instead of the blocks $\Sigma_\ell = \{(A_i \preccurlyeq B_i) \mid k_{\ell-1} \leq i < k_\ell\}$. $\square$

The proof of Theorem 5.5.4 is also readily adapted to cover most extensions of $\mathbb{V}_{\preccurlyeq}$ as well.

Corollary 5.5.5. *The logics $\mathbb{VN}_{\preccurlyeq}, \mathbb{VW}_{\preccurlyeq}, \mathbb{VC}_{\preccurlyeq}, \mathbb{VA}_{\preccurlyeq}$ and $\mathbb{VNA}_{\preccurlyeq}$ have the Craig interpolation property.*

Proof. For $\mathbb{VN}_{\preccurlyeq}$ we need to consider the additional case of a rule $\mathsf{R}_{m,0}$. Similar to the case of $n = 1$ in the procedure given above we replace the missing interpolants I_1 and J_1 by $\bot$ and J_2 by the interpolant for the premiss $\Rightarrow B_1, \ldots, B_n$. For the rules $\mathsf{W}_{m,n}$ of $\mathcal{R}_{\mathbb{VW}_{\preccurlyeq}}$ we modify the construction by replacing the interpolants J_1, J_2 with the interpolant J of the contextual premiss resp. its negation. The case that $n = 0$ does not cause any problems. For the rule R_{W2} we simply use the interpolant of the contextual premiss. For R_{C2}, if the splitting of the conclusion is $\Gamma_1 \mid \Gamma_2, (A \preccurlyeq B) \Rightarrow \Delta_1 \mid \Delta_2$ and the interpolants of the corresponding splittings of the premisses are I_1 and I_2 we use $I := I_1 \wedge I_2$. In case the splitting of the conclusion is $\Gamma_1, (A \preccurlyeq B) \mid \Gamma_2 \Rightarrow \Delta_1 \mid \Delta_2$, then we use $I_1 \vee I_2$. In each case it is easy to check that the formula I is an interpolant. Finally, for the rules $\mathsf{A}_{m,n}$ of $\mathcal{R}_{\mathbb{VA}_{\preccurlyeq}}$ and $\mathcal{R}_{\mathbb{VNA}_{\preccurlyeq}}$ the proof carries over almost verbatim, we only need to add the context to the premisses. □

Remark 5.5.6. The problem with adapting this proof to the case of $\mathbb{VT}_{\preccurlyeq}$ is that in order to prove Claim 2 for $S = \emptyset$ we would need to apply rule T_m. For this the second argument of one of the V_k would need to contain the interpolant for the contextual premiss. But then showing Claim 1, in particular the case $\Gamma_1, W_k \Rightarrow \Delta_1, V_k$ becomes problematic, since then the second argument of V_k does not entail the disjunction of the B_i anymore. It is not clear whether this problem can be avoided.

Since the weak and the strong counterfactual implication can be defined in terms of the comparative plausibility operator, the interpolation results can also easily be transferred to the logics formulated using these connectives.

Corollary 5.5.7. *Let $* \in \{\Box\!\!\Rightarrow, \Box\!\!\rightarrow\}$. Then the logics $\mathbb{V}_*, \mathbb{VN}_*, \mathbb{VW}_*, \mathbb{VC}_*, \mathbb{VA}_*$ and $\mathbb{VNA}_*$ have the Craig interpolation property.*

Proof. Let $*$ be one of $\Box\!\!\Rightarrow, \Box\!\!\rightarrow$ and let $\mathcal{L}$ be one of the logics specified. Furthermore, write $\tau : \mathcal{F}(*) \to \mathcal{F}(\preccurlyeq)$ and $\sigma : \mathcal{F}(\preccurlyeq) \to \mathcal{F}(*)$ for the translation functions given by extending the translations

$$(p \Box\!\!\Rightarrow q) \to \neg((p \wedge \neg q) \preccurlyeq (p \wedge q)) \quad \text{resp.} \quad (p \Box\!\!\rightarrow q) \to (\bot \preccurlyeq p) \vee \neg((p \wedge \neg q) \preccurlyeq (p \wedge q))$$

and

$$(p \preccurlyeq q) \to \neg(p \vee q \Box\!\!\Rightarrow \neg p) \quad \text{resp.} \quad (p \preccurlyeq q) \to (p \vee q \Box\!\!\rightarrow \neg(p \vee q)) \vee \neg(p \vee q \Box\!\!\rightarrow \neg p)$$

from Table 5.1 on p. 144 in the obvious way. Then it is not too difficult to check that for every formula $A \in \mathcal{F}(*)$ we have $\models_{\mathcal{L}_*} A \leftrightarrow \sigma \circ \tau(A)$. Thus for a $\mathcal{L}_*$-valid implication $A \to B$ we construct an interpolant I using the interpolant I' for the $\mathcal{L}_{\preccurlyeq}$-valid implication $\tau(A) \to \tau(B)$

by setting $I := \sigma(I')$. Then we have:

$$\begin{aligned} \models_{\mathcal{L}_*} A \to B &\iff \models_{\mathcal{L}_\preccurlyeq} \tau(A) \to \tau(B) \\ &\implies \models_{\mathcal{L}_\preccurlyeq} \tau(A) \to I \text{ and } \models_{\mathcal{L}_\preccurlyeq} I \to \tau(B) \\ &\implies \models_{\mathcal{L}_*} \sigma \circ \tau(A) \to \sigma(I) \text{ and } \models_{\mathcal{L}_*} \sigma(I) \to \sigma \circ \tau(B) \\ &\iff \models_{\mathcal{L}_*} A \to \sigma(I) \text{ and } \models_{\mathcal{L}_*} \sigma(I) \to B \end{aligned}$$

Also it is clear that $\mathsf{var}(\sigma(I)) \subseteq \mathsf{var}(A) \cap \mathsf{var}(B)$. Thus I indeed is an interpolant for the implication $A \to B$. □

5.6 Notes

Conditional Logics The formulation of conditional logics in terms of sphere semantics was introduced in [Lew73b, Lew73a]. In particular [Lew73a] contains a thorough discussion of the philosophical motivation and can only be recommended. A slightly earlier formulation of counterfactual implications in terms of closest or most similar worlds can be found in [Sta68] with many of the technical details in [ST70]. Good overviews over the general problem and the history of conditional logics are also given in [NC01] and [AC09]. The article [GGOS09] includes a good and compact introduction to the various semantics. Due to the plethora of proposed systems for conditional logics and the often slightly different axiomatisations by different authors the task of navigating the conditional logic landscape can be a bit daunting. The article [Nej91] provides welcome comparisons between the different formulations. The original semantically driven proofs for the complexity results considered in this chapter can be found in [FH94]. They are obtained using small model theorems and also give rise to complexity results for fragments of the logics obtained by bounding the modal nesting depth. Finally it should be noted that decidability for most of the systems of conditional logic also follows from the generic decidability results for logics axiomatised by non-iterative axioms provided in [Lew74].

Sequent calculi for conditional logics. The constructed calculi for strong systems of conditional logics along with the results about interpolation were presented in [LP12b]. In the case of $\mathbb{VC}_\preccurlyeq$, our construction of a principal-cut and contraction closed rule set yielded the sequent analogue of the tableau calculus for the same logic introduced in [dS83] and subsequently corrected in [Gen92]. These works also presented a calculus for the logic $\mathbb{VCS}_\preccurlyeq$, an extension of the logic $\mathbb{VC}_\preccurlyeq$ with *Stalnaker's Axiom* (S) $(p \land q \preccurlyeq p \land \neg q) \land (p \land \neg q \preccurlyeq p \land q) \to (\bot \preccurlyeq p)$. Furthermore, these systems were used to give a decision procedure for the logics $\mathbb{VC}_\preccurlyeq$ and $\mathbb{VCS}_\preccurlyeq$. Most of the other approaches towards proof theory for conditional logic concentrate on the formulation with the weak counterfactual implication as the main connective. Labelled tableau calculi for conditional logics in the language with $\Box\!\!\rightarrow$ extending the slightly weaker logic PCL including the here considered logics $\mathbb{V}_{\Box\!\rightarrow}$, $\mathbb{VW}_{\Box\!\rightarrow}$, $\mathbb{VC}_{\Box\!\rightarrow}$ and extensions with the axiom (CEM) $(p \Box\!\!\rightarrow q) \lor (p \Box\!\!\rightarrow \neg q)$ or the *universality axioms* are considered in [GGOS03, GGOS09]. Since in these calculi each additional semantic property corresponds to a single rule, they have the very desirable property of being *modular*. On the other hand they make essential use of labels and only yield decision procedures of slightly sub-optimal NExptime-complexity. This work also contains a very readable discussion of other approaches towards calculi for these logics. An unlabelled system for the weakest of the logics considered there, the logic PCL, was also given in [SPH10], where it is called $\mathcal{S}$. The calculus gives rise to a Pspace-decision procedure, but is slightly complicated.

Interpolation No interpolation results for stronger systems of conditional logic seem to have been established before our [LP12b]. A stronger form of the Craig interpolation property, the *uniform interpolation property*, states that the interpolant for derivable implications $A \to B$ is the same for all formulae B with the same common variables with A. The uniform interpolation property for the logics $\mathbb{V}$ and $\mathbb{VN}$ follows from recent results in [Pat13] where using coalgebraic semantics the very general result that every logic axiomatised in rank-1 enjoys the uniform interpolation property is established.

6 Conclusion

6.1 Summary

Let us briefly summarise what we have achieved in this work. Motivated by the desire to develop generic proof theoretic methods for the treatment of possibly non-normal modal logics which might be based on non-classical propositional logics we found that we needed to specify a format of sequent rules first. This led to the introduction of the formats of rules with context restrictions and as a weaker version shallow rules, both extensions of the earlier considered format of one-step rules. A closer investigation of sequent calculi given by sets of rules with restrictions yielded a generic cut elimination theorem for rule sets satisfying the criteria of principal-cut, mixed-cut, context-cut and contraction closure, which can be checked by considering single rules or pairs of rules from the rule set. For rule sets which in addition are tractable we were able to show a generic EXPTIME-decidability result and noted that the complexity can be lowered to (optimal) PSPACE in case all the rules are shallow.

We then investigated the connections between rules with context restrictions and axioms for Hilbert-style proof systems and developed a syntactic characterisation of a class of axioms, the translatable clauses, which very closely corresponds to the class of rules with context restrictions. This correspondence was established by explicit and automatic translations between axioms and rules making use of the concepts of ω-sets of axioms and proto rules. Limiting the rule format to shallow rules resulted in a correspondence between shallow rules and non-iterative translatable clauses, which for modal logics based on classical propositional logic gives a correspondence between shallow rules and non-iterative axioms in general. We then used these correspondences to derive a number of results stating amongst others that modal logic T cannot be captured by one-step rules, that $\mathsf{K4}$ cannot be captured by shallow rules and that the logics B, GL and $\mathsf{S5}$ cannot be captured by mixed-cut closed sets of rules with context restrictions.

While the translation from rules to axioms gave us mainly negative results the translation from axioms to rules provided the starting point for an automatic construction of sequent calculi from finite sets of Hilbert axioms. In order to produce cut-free sequent calculi we investigated the concept of the cut closure of a rule set and found a representation of the rules in the cut closure in terms of cut trees. Closing the original rule set under cuts with small rules in a preprocessing step allowed us to limit the size and depth of the necessary cut trees and thus to show that the cut closure of a finite set of rules is tractable. Using this we saw that contraction closure of the cut closure of finite sets of shallow rules entails PSPACE-decidability of the corresponding logic. For cases where the cut closure of a rule set cannot be seen to be contraction closed or where for other reasons the manual construction of a principal-cut closed set of rules with context restrictions is desired we introduced a graphical representation of sequents as doodles and sequent rules as rule doodles. This tool allows us to manually perform the operations of cuts between rules and contractions of rules in a very intuitive way by simply connecting heads and tails of doodles. As simple examples for this technique we considered Elgesem's logic of agency and ability and weaker systems of conditional logic.

Finally, we applied the earlier introduced techniques and considered the extended example of strong systems of conditional logic. The main results here were the construction of cut-free sequent calculi for Lewis' conditional logic $\mathbb{V}$ and extensions in the language of the comparative plausibility operator $\preccurlyeq$ using rule doodles and the adaption of these calculi to the strong counterfactual implication $\Box\!\!\Rightarrow$ as the main connective. In the case of $\mathbb{VC}$ we recovered the sequent analogue of a previously known tableau calculus in a systematic and purely syntactic way. Using the generic decision procedures developed earlier these systems yielded purely syntactic decision procedures which, except for the logics $\mathbb{VA}$ and $\mathbb{VNA}$, are complexity optimal. As a further application we used the calculi to derive interpolation results for all of the logics under consideration apart from the logic $\mathbb{VT}$.

6.2 Applications and Alternative Approaches

It is to be hoped that the methods and results developed in this work will provide a toolkit for the proof-theoretical treatment of non-classical logics in the spirit of logic engineering. Ideally they should on the one hand enable a researcher interested in a particular logic to evaluate whether this logic can be captured proof-theoretically in the framework of standard sequent calculi. The methods for proving the limitative results of Section 3.4 provide some tools for this purpose. On the other hand, if the logic at hand can be captured in the framework of standard sequent calculi, then the main difficulty lies in the construction of a cut-free sequent calculus for this logic. Here the generic cut elimination result of Section 2.4 provides a guidance towards which kinds of logical rules should be constructed. Moreover, the translation results of Chapter 3 together with the results about saturating a rule set under cuts using the graphical representation of sequents as doodles and sequent rules as rule doodles provide some tools to actually construct a calculus with logical rules satisfying the criteria for cut elimination. Of course ultimately we would like to provide a fully automatic construction of cut-free sequent calculi and (perhaps more importantly) decision procedures for suitable modal logics given as a set of Hilbert-style axioms. While the tools presented here do not yet give rise to such a fully automatic construction, they can be used to identify the crucial property of contraction-closure of the rule set, and thus for some logics give rise to proofs of decidability and complexity results via establishing this property for a set of sequent rules.

The methods and results presented here are mainly theoretical, but in particular from the engineering point of view it would also be desirable to have an implementation of these results which given a number of axioms for a logic produces a cut-free sequent calculus or preferably even a decision procedure for this logic. A similar implementation based on a different theory exists in the Paralyzer system [CLSZ13] which transforms axioms of a certain form into rules of a cut-free sequent calculus based on positive propositional logic. Such an implementation would be especially powerful when combined with an implementation of the generic decision procedures from Section 2.7 similar to the CoLoSS system [CMPS09] which implements a generic decision procedure for rank-1 modal logics. If these methods are to be used as tools in this general sense, it would of course be important to have a very high degree of confidence in both the theoretical results and the implementations. On the implementation side for this purpose it might prove beneficial to make use of e.g. a programming language supporting functional programming with dependent types and its type checking feature. Since correct programs in such a language need to pass type checking we would have a higher confidence in the correct implementation of the compilations. On the theoretical side on the other hand this could be ensured by using proof assistants such as Coq or Isabelle to formalise and verify the main results of the theory, in particular the generic cut elimination theorem. Since the proof of this theorem relies on the analysis of a variety of different cases it would lend itself to such a formalisation.

Depending on which kinds of logics we are interested in and which applications we have in mind we might also consider frameworks different from that of standard sequents considered in this work. Thus for example if we are interested in modal logics based only on classical propositional logic we might modify the theory presented here and develop it on the basis of *one-sided sequents* in the style of Schütte-Tait [TS00] instead of two-sided sequents. This would arguably lead to a more streamlined and elegant formulation of the theory, in particular if for each modality we also have a dual modality allowing us to push negations all the way inside the formulae. An approach particularly suitable for dealing with modal logics based on intuitionistic propositional logic on the other hand would be to consider the more general framework of *hypersequents* [Avr96] instead of sequents. This approach has already been employed successfully e.g. in the treatment of substructural logics [CGT08, CGT12] and intermediate logics [CMS13]. It is to be expected that the additional structure provided by hypersequents would allow for a treatment of some occurrences of non-invertible connectives in the axioms, in particular of disjunctions. On the other hand we might lose some of the nice properties and applications provided by the standard sequent framework: decision procedures based on hypersequents in general seem to be of higher complexity compared to those based on standard sequents, and the hypersequent framework seems not to facilitate proofs of the interpolation property. We might also be willing to trade in even more of the tractability and applications of the standard sequent framework for the even further increased expressivity allowed by the framework of *display calculi* [Bel82]. This framework is particularly convenient for logics with pairs of connectives satisfying some kind of residuation property such as the forward and backward looking modalities of temporal modal logic. The results from [Kra96] and [CR13] on the connection between axioms and structural rules in this framework would provide a benchmark for the theory and moreover a very interesting opportunity to merge different approaches.

On the other hand, if the considered modalities are normal and in particular if we have a Kripke-style semantics, we could also move to more semantically motivated frameworks such as those of *nested sequents* [Brü09] or *tree-hypersequents* [Pog09, Pog11]. Since in these frameworks the tree-like structure of (unravellings of) Kripke frames is matched by the tree-like structure of nested sequents resp. tree-hypersequents, they might be better suited for the construction of sequent calculi from semantical properties, although no generic treatment seems to have been developed yet. Finally, when starting from such semantical characterisations of a logic we might also consider internalising the Kripke semantics in the sequent framework by moving to *labelled sequents*. This would also allow us to build on and combine our methods with the extensive theory for this kind of sequent calculi developed in [NvP01, Neg05].

6.3 The Future

Concerning further research apart from the above mentioned modifications to incorporate different frameworks there are two linked albeit slightly different major directions. From the point of view of *classification* it would be very interesting to extend the limitative results of Section 3.4 to formally establish the need for additional machinery beyond the standard sequent framework for certain modal logics. From a more *constructive* point of view the further development of the generic methods for the construction of cut-free sequent calculi is well worth investigating. The most pressing concrete problem from the latter point of view is the hole in the construction of contraction closed rule sets for non-iterative modal logics via cut trees. We conjecture that it should be possible to close this hole by establishing a polynomial upper bound on the number of duplicate formulae needed to derive a sequent, but a proof for this has eluded all our efforts so far.

Problem 6.3.1. *Extend the method of cut trees to produce also contraction closed rule sets.*

From the classification point of view the most promising concrete problem is that of extending the methods for showing limitative results by relaxing the condition of mixed-cut closure of the rule set. While the relatively general form of context restrictions might make it difficult to restrict the precise form of the corresponding axioms, it might still be possible to exploit the fact that context formulae do not share variables with each other or with the principal formulae.

Problem 6.3.2. *Show limitative results for rules with restrictions in general.*

As a more general direction of further research it would be very interesting to extend the format of rules with context restrictions to include more logics such as B or GL. A promising way to do this seems to be the extension of context restrictions towards the notion of context relations from [AL11]. An extension of the correspondence results of Chapter 3 to such rules would serve to further classify modal logics and to better understand the strengths and limitations of the different variants of sequent calculi. As a long term goal the resulting theory should also incorporate extensions of the sequent format such as hypersequents, nested or labelled sequents.

Research Programme 6.3.3. *Develop an extensive classification of modal logics according to the strength of logical rules necessary to capture them in a cut-free sequent calculus.*

Finally, concerning the example of conditional logics, since the weak counterfactual implication is usually taken to be the main connective, it would be very nice to have explicit formulations of the sequent calculi presented in terms of this connective.

Problem 6.3.4. *Find simple standard cut-free sequent calculi for the strong systems of conditional logic extending* $\mathbb{V}$ *in the language of the weak counterfactual implication* $\Box\!\!\rightarrow$.

Bibliography

[AB09] Sanjeev Arora and Boaz Barak. *Computational Complexity - A Modern Approach.* Cambridge University Press, 2009.

[AC09] Horacio Arlo-Costa. The logic of conditionals. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* http://plato.stanford.edu/archives/spr2009/entries/logic-conditionals/, spring 2009 edition, 2009.

[AL01] Arnon Avron and Iddo Lev. Canonical propositional Gentzen-type systems. In Rajeev Goré, Alexander Leitsch, and Tobias Nipkow, editors, *IJCAR 2001*, volume 2083 of *LNAI*, pages 529–544. Springer-Verlag Berlin Heidelberg, 2001.

[AL09] Arnon Avron and Ori Lahav. Canonical constructive systems. In Martin Giese and Arild Waaler, editors, *TABLEAUX 2009*, volume 5607 of *LNAI*, pages 62–76. Springer-Verlag Berlin Heidelberg, 2009.

[AL11] Arnon Avron and Ori Lahav. Kripke semantics for basic sequent systems. In Kai Brünnler and George Metcalfe, editors, *TABLEAUX 2011*, volume 6793 of *LNAI*, pages 43–57. Springer-Verlag Berlin Heidelberg, 2011.

[AOP12] Régis Alenda, Nicola Olivetti, and Gian Luca Pozzato. Nested sequent calculi for conditional logics. In Luis Fari nas del Cerro, Andreas Herzig, and Jérôme Mengin, editors, *JELIA 2012*, volume 7519 of *LNCS*, pages 14–27. Springer-Verlag Berlin Heidelberg, 2012.

[Arn08] Michael Arndt. *Logical Tomography.* PhD thesis, Eberhardt-Karls-Universität Tübingen, 2008.

[Avr96] Arnon Avron. The method of hypersequents in the proof theory of propositional non-classical logics. In Wilfrid Hodges, Martin Hyland, Charles Steinhorn, and John Truss, editors, *Logic: From Foundations to Applications.* Clarendon Press New York, 1996.

[AZ08] Arnon Avron and Anna Zamansky. Canonical calculi with (n,k)-ary quantifiers. *Log. Methods Comput. Sci.*, 4(3:2):1–23, 2008.

[BCar] Matthias Baaz and Agata Ciabattoni. Proof theory of witnessed Gödel logic: a negative result. *J. Log. Comput.*, (to appear).

[BdPR01] Gianluigi Bellin, Valeria de Paiva, and Eike Ritter. Extended Curry-Howard correspondence for a basic contructive modal logic. In *Methods for Modalities II*, 2001.

[BdRV01] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*. Cambridge University Press, 2001.

[Bel82] Nuel D. Belnap. Display logic. *J. Philos. Logic*, 11:375–417, 1982.

[BG13] Nick Bezhanishvili and Silvio Ghilardi. Bounded proofs and step frames. In Didier Galmiche and Dominique Larchey-Wendling, editors, *TABLEAUX 2013*, volume 8123 of *LNCS*, pages 44–58. Springer-Verlag Berlin Heidelberg, 2013.

[BLZ12] Matthias Baaz, Ori Lahav, and Anna Zamansky. Effective finite-valued semantics for labelled calculi. In Bernhard Gramlich, Dale Miller, and Ulrike Sattler, editors, *IJCAR 2012*, volume 7364 of *LNCS*, pages 52–66, 2012.

[Brü09] Kai Brünnler. Deep sequent systems for modal logic. *Arch. Math. Log.*, 48:551–577, 2009.

[Bur81] John P. Burgess. Quick completeness proofs for some logics of conditionals. *Notre Dame J. Form. Log.*, 22(1):76—84, Jan 1981.

[CGT08] Agata Ciabattoni, Nikolaos Galatos, and Kazushige Terui. From axioms to analytic rules in nonclassical logics. In *LICS*, pages 229–240. IEEE Computer Society, 2008.

[CGT12] Agata Ciabattoni, Nikolaos Galatos, and Kazushige Terui. Algebraic proof theory for substructural logics: Cut-elimination and completions. *Ann. Pure Appl. Logic*, 163:266–290, 2012.

[Che75] Brian F. Chellas. Basic conditional logic. *J. Philos. Logic*, 4:133–153, Jan 1975.

[Che80] Brian F. Chellas. *Modal Logic*. Cambridge University Press, 1980.

[CKS81] Ashok K Chandra, Dexter C Kozen, and Larry J Stockmeyer. Alternation. *J. Assoc. Comput. Mach.*, 28(1):114–133, 1981.

[CLSZ13] Agata Ciabattoni, Ori Lahav, Lara Spendier, and Anna Zamansky. Automated support for the investigation of paraconsistent and other logics. In Sergei Artemov and Anil Nerode, editors, *LFCS 2013*, volume 7734 of *LNCS*, pages 119–133. Springer-Verlag Berlin Heidelberg, 2013.

[CMPS09] Georgel Calin, Robert S. R. Myers, Dirk Pattinson, and Lutz Schröder. CoLoSS: The coalgebraic logic satisfiability solver. *Electr. Notes Theor. Comput. Sci.*, 231:41–54, Jan 2009.

[CMS13] Agata Ciabattoni, Paolo Maffezioli, and Lara Spendier. Hypersequent and labelled calculi for intermediate logics. In Didier Galmiche and Dominique Larchey-Wendling, editors, *TABLEAUX 2013*, volume 8123 of *LNCS*, pages 81–96. Springer-Verlag Berlin Heidelberg, 2013.

[Coe10] Bob Coecke. Quantum picturalism. *Contemporary Physics*, 51(1):59–83, 2010.

[CR13] Agata Ciabattoni and Revantha Ramanayake. Structural extensions of display calculi: a general recipe. In Leonid Libkin, Ulrich Kohlenbach, and Ruy de Queiroz, editors, *WoLLIC 2013*, volume 8071 of *LNCS*, pages 81–95. Springer-Verlag Berlin Heidelberg, 2013.

[Cro01] Tristan Crolard. Subtractive logic. *Theoret. Comput. Sci.*, 254:151–185, 2001.

[CST09] Agata Ciabattoni, Lutz Straßburger, and Kazushige Terui. Expanding the realm of systematic proof theory. In Erich Grädel and Reinhard Kahle, editors, *CSL 2009*, volume 5771 of *LNCS*, pages 163–178. Springer-Verlag Berlin Heidelberg, 2009.

[CT06a] Agata Ciabattoni and Kazushige Terui. Modular cut-elimination: Finding proofs or counterexamples. In Miki Hermann and Andrei Voronkov, editors, *LPAR 2006*, volume 4246 of *LNAI*, pages 135–149. Springer-Verlag Berlin Heidelberg, 2006.

[CT06b] Agata Ciabattoni and Kazushige Terui. Towards a semantic characterization of cut-elimination. *Studia Logica*, 82:95–119, 2006.

[CZ97] Alexander Chagrov and Michael Zakharyaschev. *Modal Logic.* Oxford University Press, 1997.

[Die06] Reinhard Diestel. *Graph Theory.* Springer-Verlag Berlin Heidelberg, 3rd edition, 2006.

[dS83] Harrie C.M. de Swart. A Gentzen- or Beth-type system, a practical decision procedure and a constructive completeness proof for the counterfactual logics VC and VCS. *J. Symb. Log.*, 48(1):1–20, 1983.

[EFT94] Heinz-Dieter Ebbinghaus, Jörg Flum, and Wolfgang Thomas. *Mathematical Logic.* Springer-Verlag New York, New York, 2 edition, 1994.

[EFT96] Heinz-Dieter Ebbinghaus, Jörg Flum, and Wolfgang Thomas. *Einführung in die mathematische Logik.* Spektrum, Akad. Verl., Heidelberg, Berlin, 4 edition, 1996.

[Elg97] Dag Elgesem. The modal logic of agency. *Nordic Journal of Philosophical Logic*, 2(2):1–46, 1997.

[FH94] Nir Friedman and Joseph Y. Halpern. On the complexity of conditional logics. In Jon Doyle, Erik Sandewall, and Pietro Torasso, editors, *KR'94*, pages 202–213. Morgan Kaufmann, 1994.

[FM97] M. Fairtlough and M.V. Mendler. Propositional lax logic. *Inform. and Comput.*, 137(1):1–33, 1997.

[Gen34] Gerhard Gentzen. Untersuchungen über das logische Schließen. I. *Math. Z.*, 39(2):176–210, 1934.

[Gen35] Gerhard Gentzen. Untersuchungen über das logische Schließen. II. *Math. Z.*, 39(3):405–431, 1935.

[Gen92] Ian P. Gent. A sequent- or tableau-style system for Lewis's counterfactual logic VC. *Notre Dame J. Form. Log.*, 33(3):369–382, 1992.

[GG08] Alessio Guglielmi and Tom Gundersen. Normalisation control in deep inference via atomic flows. *Log. Methods Comput. Sci.*, 4(1:9):1–36, 2008.

[GGGP11] Valerio Genovese, Laura Giordano, Valentina Gliozzi, and Gian Luca Pozzato. A conditional constructive logic for access control and its sequent calculus. In Kai Brünnler and George Metcalfe, editors, *TABLEAUX 2011*, volume 6793 of *LNAI*, pages 164–179. Springer-Verlag Berlin Heidelberg, 2011.

[GGN12] Deepark Garg, Valerio Genovese, and Sara Negri. Countermodels from sequent calculi in multi-modal logics. In *LICS 2012*, pages 315–324, 2012.

[GGOS03] Laura Giordano, Valentina Gliozzi, Nicola Olivetti, and Camilla Schwind. Tableau calculi for preference-based conditional logics. In Marta Cialdea Mayer and Fiora Pirri, editors, *TABLEAUX 2003*, volume 2796 of *LNCS*, pages 81–101. Springer-Verlag Berlin Heidelberg, 2003.

[GGOS09] Laura Giordano, Valentina Gliozzi, Nicola Olivetti, and Camilla Schwind. Tableau calculus for preference-based conditional logics: PCL and its extensions. *ACM Trans. Comp. Log.*, 10(3):21:1–21:50, 2009.

[Ghi99] Silvio Ghilardi. Unification in intuitionistic logic. *J. Symb. Log.*, 64(2):859–880, 1999.

[Gir87] Jean-Yves Girard. Linear logic. *Theoret. Comput. Sci.*, 50:1–102, 1987.

[Gob74] Louis F. Goble. Gentzen systems for modal logic. *Notre Dame J. Form. Log.*, 15(3):455–461, 1974.

[GR05] Guido Governatori and Antonino Rotolo. On the axiomatisation of Elgesem's logic of agency and ability. *J. Philos. Logic*, 34(4):403–431, 2005.

[GR08] Rajeev Goré and Revantha Ramanayake. Valentini's cut-elimination for provability logic resolved. In Carlos Areces and Robert Goldblatt, editors, *AiML 7*, pages 67–86. College Publications, 2008.

[HC96] George Edward Hughes and Max J. Cresswell. *A New Introduction to Modal Logic*. Routledge, London, 1996.

[Her22] Paul Hertz. Über Axiomensysteme für beliebige Satzsysteme I. Sätze ersten Grades. *Math. Ann.*, 87(3-4):246–269, 1922.

[Her23] Paul Hertz. Über Axiomensysteme für beliebige Satzsysteme II. Sätze höheren Grades. *Math. Ann.*, 89(1-2):76–102, 1923.

[Her29] Paul Hertz. Über Axiomensysteme für beliebige Satzsysteme. *Math. Ann.*, 101(1):457–514, 1929.

[HSZ96] Alain Heuerding, Michael Seyfried, and Heinrich Zimmermann. Efficient loop-check for backward proof search in some non-classical propositional logics. In Pierangelo Miglioli, Ugo Moscato, Daniele Mundici, and Mario Ornaghi, editors, *TABLEAUX 1996*, volume 1071 of *LNCS*, pages 210–225. Springer-Verlag Berlin Heidelberg, 1996.

[IM09] Rosalie Iemhoff and George Metcalfe. Proof theory for admissible rules. *Ann. Pure Appl. Logic*, 159(1-2):171–186, 2009.

[ISH65] Philip M. Lewis II, Richard Edwin Stearns, and Juris Hartmanis. Memory bounds for recognition of context-free and context-sensitive languages. In *SWCT (FOCS) 1965*, pages 191–202, 1965.

[Joh37] Ingebrigt Johansson. Der Minimalkalkül, ein reduzierter intuitionistischer Formalismus. *Compos. Math.*, 4:119–136, 1937.

[Kle52] Stephen Cole Kleene. *Introduction to Metamathematics*. North-Holland, Amsterdam, 1952.

[KNSS95] Ágnes Kurucz, István Németi, Ildikó Sain, and András Simon. Decidable and undecidable logics with a binary modality. *J. Log. Lang. Inf.*, 4:191–206, 1995.

[Kra96] Marcus Kracht. Power and weakness of the modal display calculus. In Heinrich Wansing, editor, *Proof Theory of Modal Logic*, pages 93–121. Kluwer, 1996.

[Lad77] Richard E. Ladner. The computational complexity of provability in systems of modal logic. *SIAM J. Comput.*, 6(3):467–480, 1977.

[Lah13] Ori Lahav. From frame properties to hypersequent rules in modal logics. In *LICS 2013*, 2013.

[Lei81] Daniel Leivant. On the proof theory of the modal logic for arithmetic provability. *J. Symb. Log.*, 46(3):531–538, 1981.

[Lew73a] David Lewis. *Counterfactuals*. Blackwell, 1973.

[Lew73b] David Lewis. Counterfactuals and comparative possibility. *J. Philos. Logic*, 2:418–446, 1973.

[Lew74] David Lewis. Intensional logics without iterative axioms. *J. Phil. Log.*, 3:457–466, 1974.

[LP11] Björn Lellmann and Dirk Pattinson. Cut elimination for shallow modal logics. In Kai Brünnler and George Metcalfe, editors, *TABLEAUX 2011*, volume 6793 of *LNAI*, pages 211–225. Springer-Verlag Berlin Heidelberg, 2011.

[LP12a] Björn Lellmann and Dirk Pattinson. Graphical construction of cut-free sequent systems suitable for backwards proof search (extended abstract). In Renate A. Schmidt and Fabio Papacchini, editors, *Proceedings of the 19th Automated Reasoning Workshop*, pages 17–18, Manchester, 2012. School of Computer Science, The University of Manchester.

[LP12b] Björn Lellmann and Dirk Pattinson. Sequent systems for Lewis' conditional logics. In Luis Fariñas del Cerro, Andreas Herzig, and Jerome Mengin, editors, *JELIA 2012*, volume 7519 of *LNCS*, pages 320–332. Springer-Verlag Berlin Heidelberg, 2012.

[LP13a] Björn Lellmann and Dirk Pattinson. Constructing cut free sequent systems with context restrictions based on classical or intuitionistic logic. In Kamal Lodaya, editor, *ICLA 2013*, volume 7750 of *LNAI*, pages 148–160. Springer-Verlag Berlin Heidelberg, 2013.

[LP13b] Björn Lellmann and Dirk Pattinson. Correspondence between modal Hilbert axioms and sequent rules with an application to S5. In Didier Galmiche and Dominique Larchey-Wendling, editors, *TABLEAUX 2013*, volume 8123 of *LNCS*, pages 219–233. Springer-Verlag Berlin Heidelberg, 2013.

[Moe01] A. Moen. The proposed algorithms for eliminating cuts in the provability calculus GLS do not terminate. Technical report, Norwegian Computing Center, 2001.

[NC01] Donald Nute and Charles B. Cross. Conditional logic. In Dov M. Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic*, volume 4, pages 1–98. Kluwer, 2001.

[Neg03] Sara Negri. Contraction-free sequent calculi for geometric theories with an application to Barr's theorem. *Arch. Math. Log. (AML)*, 42(4):389–401, 2003.

[Neg05] Sara Negri. Proof analysis in modal logic. *J. Philos. Logic*, 34:507–544, 2005.

[Nej91] Wolfgang Nejdl. The p-systems: A systematic classification of logics of nonmonotonicity. In *AAAI*, pages 366 – 372, 1991.

[Nut80] Donald Nute. *Topics in Conditional Logic*. Reidel, Dordrecht, 1980.

[NvP98] Sara Negri and Jan von Plato. Cut elimination in the presence of axioms. *Bull. Symbolic Logic*, 4(4):418–435, 1998.

[NvP01] Sara Negri and Jan von Plato. *Structural proof theory*. Cambridge University Press, 2001.

[OPS07] Nicola Olivetti, Gian Luca Pozzato, and Camilla Schwind. A sequent calculus and a theorem prover for standard conditional logics. *ACM Trans. Comput. Log.*, 8(4):22/1–22/51, 2007.

[OS01] Nicola Olivetti and Camilla Schwind. A calculus and complexity bound for minimal conditional logic. In Antonio Restivo, Simona Ronchi Della Rocca, and Luca Roversi, editors, *ICTCS 2001*, volume 2202 of *LNCS*, pages 384–404. Springer-Verlag Berlin Heidelberg, 2001.

[Pap94] Christos H. Papadimitriou. *Computational Complexity*. Addison-Wesley, 2 edition, 1994.

[Pat13] Dirk Pattinson. The logic of exact covers: Completeness and uniform interpolation. In *LICS 2013*, 2013.

[Pog08] Francesca Poggiolesi. A cut-free simple sequent calculus for modal logic S5. *Rev. Symb. Log.*, 1(1):3–15, 2008.

[Pog09] Francesca Poggiolesi. The method of tree-hypersequents for modal propositional logic. In David Makinson, Jacek Malinkowski, and Heinrich Wansing, editors, *Towards Mathematical Philosophy*, volume 28 of *Trends In Logic*, pages 31–51. Springer Science + Business Media B.V., 2009.

[Pog11] Francesca Poggiolesi. *Gentzen Calculi for Modal Propositional Logic*, volume 32 of *Trends In Logic*. Springer-Verlag Berlin Heidelberg, 2011.

[Poz10] Gian Luca Pozzato. *Conditional and Preferential Logics: Proof Methods and Theorem Proving*, volume 208 of *Frontiers in Artificial Intelligence and Applications*. IOS Press Amsterdam, 2010.

[PS08] Dirk Pattinson and Lutz Schröder. Admissibility of cut in coalgebraic logics. *Electr. Notes Theor. Comput. Sci.*, 203(5):221–241, 2008.

[PS09] Dirk Pattinson and Lutz Schröder. Generic modal cut elimination applied to conditional logics. In Martin Giese and Arild Waaler, editors, *TABLEAUX 2009*, volume 5607 of *Lecture Notes in Computer Science*, pages 280–294. Springer-Verlag Berlin Heidelberg, 2009.

[PS10] Dirk Pattinson and Lutz Schröder. Cut elimination in coalgebraic logics. *Inf. Comput.*, 208(12):1447–1468, 2010.

[PS11] Dirk Pattinson and Lutz Schröder. Generic modal cut elimination applied to conditional logics. *Log. Methods Comput. Sci.*, 7(1:4):1–28, 2011.

[Ras07] João Rasga. Sufficient conditions for cut elimination with complexity analysis. *Ann. Pure Appl. Logic*, 149:81–99, 2007.

[Res00] Greg Restall. *An Introduction to Substructural Logics*. Routledge, 2000.

[Sch07] Lutz Schröder. A finite model construction for coalgebraic modal logic. *J. Log. Algebr. Program.*, 73(1-2):97–110, 2007.

[SP08] Lutz Schröder and Dirk Pattinson. Shallow models for non-iterative modal logics. In Andreas R. Dengel, Karsten Berns, Thomas M. Breuel, Frank Bomarius, and Thomas R. Roth-Berghofer, editors, *KI 2008*, volume 5243 of *LNAI*, pages 324–331. Springer-Verlag Berlin Heidelberg, 2008.

[SP09] Lutz Schröder and Dirk Pattinson. PSPACE bounds for rank-1 modal logics. *ACM Trans. Comput. Log.*, 10(2):13:1–13:33, 2009.

[SPH10] Lutz Schröder, Dirk Pattinson, and Daniel Hausmann. Optimal tableaux for conditional logics with cautious monotonicity. In Helder Coelho, Rudi Studer, and Michael Wooldridge, editors, *ECAI 2010*, volume 215 of *Frontiers in Artificial Intelligence and Applications*, pages 707–712. IOS Press, 2010.

[ST70] Robert C. Stalnaker and Richmond H. Thomason. A semantic analysis of conditional logic. *Theoria*, 36(1):23–42, 1970.

[Sta68] Robert C. Stalnaker. A theory of conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory*, volume 2 of *American Philosophical quarterly Monograph Series*. Blackwell, 1968.

[Tak92] M. Takano. Subformula property as a substitute for cut-elimination in modal propositional logics. *Mathematica Japonica*, 37:1129–1145, 1992.

[tC05] Balder ten Cate. *Model Theory for Extended Modal Languages*. PhD thesis, ILLC Amsterdam, 2005.

[Tiu06] Alwen Tiu. A system of interaction and structure II: The need for deep inference. *Log. Methods Comput. Sci.*, 2(2:4):1–24, 2006.

[TS00] Anne Sjerp Troelstra and Helmut Schwichtenberg. *Basic Proof Theory*, volume 43 of *Cambridge Tracts In Theoretical Computer Science*. Cambridge University Press, 2 edition, 2000.

[Val83] Silvio Valentini. The modal logic of provability: Cut-elimination. *J. Philos. Logic*, 12(4):471–476, 1983.

[vB76] Johan van Benthem. Modal formulas are either elementary or not $\sigma\delta$-elementary. *J. Symb. Log.*, 41(2):436–438, 1976.

[vB83] Johan van Benthem. *Modal Logic and Classical Logic*. Bibliopolis, Napoli, 1983.

[vP01] Jan von Plato. A proof of Gentzen's Hauptsatz without multicut. *Arch. Math. Log.*, 40(1):9–18, 2001.

[Wan02] Heinrich Wansing. Sequent systems for modal logics. In Dov M. Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic, Vol.8*. Springer-Verlag Berlin Heidelberg, 2002.

[Wij90] Duminda Wijesekera. Constructive modal logics I. *Ann. Pure Appl. Logic*, 50(3):271–301, 1990.

[ZA06] Anna Zamansky and Arnon Avron. Cut-elimination and quantification in canonical systems. *Studia Logica*, 82:157–176, 2006.

[ZA12] Anna Zamansky and Arnon Avron. Canonical signed calculi with multi-ary quantifiers. *Ann. Pure Appl. Logic*, 163(7):951–960, 2012.

Index