

Analytic Methods for Logics of Conditional Reasoning

DISSERTATION

submitted in partial fulfillment of the requirements for the degree of

Doktor der Technischen Wissenschaften

by

Dmitrii Rozplokhas

Registration Number 12229969

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Prof.in Dr.in Agata Ciabattoni

Second advisor: O.Univ.Prof. Dipl.-Ing. Dr.techn. Thomas Eiter

The dissertation has been reviewed by:

Nicola Olivetti

Ofer Arieli

Vienna, August 12, 2026

Dmitrii Rozplokhas

Declaration of Authorship

Dmitrii Rozplokhas

I hereby declare that I have written this thesis independently, that I have completely specified the utilized sources and resources and that I have definitely marked all parts of the work – including tables, maps and figures – which belong to other works or to the internet, literally or extracted, by referencing the source as borrowed.

I further declare that I have used generative AI tools only as an aid, and that my own intellectual and creative efforts predominate in this work. In the appendix “Overview of Generative AI Tools Used” I have listed all generative AI tools that were used in the creation of this work, and indicated where in the work they were used. If whole passages of text were used without substantial changes, I have indicated the input (prompts) I formulated and the IT application used with its product name and version number/date.

Vienna, August 12, 2026

Dmitrii Rozplokhas

Acknowledgements

Four years ago, my world was shaken by the news that I had been admitted to the outstanding LogiCS doctoral programme and have the exceptional opportunity to study logic in Vienna, its legendary citadel. These years have passed incredibly quickly, yet they have been among the happiest of my life. Along the way, I have met so many wonderful people who have shown me an unbelievable amount of kindness and given me more help than I could ever have expected. I am deeply indebted to all of them and want to thank as many as I can here, although a few pages could not possibly hold all their names or convey all the gratitude I feel.

First and foremost, I want to thank my supervisor, Agata. Thank you for the extraordinary amount of time and thought you have put into helping me grow into a logician. For the hundreds of meetings I always leave buzzing with enthusiasm and armed with suggestions whose depth I often realise only several days later. For teaching me the art of presenting, always encouraging my funny bizarre slide ideas, and invariably topping them with something much funnier and considerably more bizarre. For always looking beyond the current project, focusing both on my future and my current well-being, checking up on me and cheering me up, and providing cats when some grounding petting is needed. My scientific journey—which is hopefully only beginning—has been and will continue to be shaped by your guidance, and I cannot imagine a better guide.

Next, I want to thank those who have made my doctoral studies—and now my graduation—possible. To Nicola, who during a short two-week visit three years ago converted me once and for all to looking at proof theory through the lens of semantics—a perspective that underlies this whole thesis. To Ofer, who was a kind and encouraging session chair for my very first talk at KR and then climbed the hills of Rhodes with me afterwards. Thank you both for taking on the review of this thesis, for doing so with such efficiency, and for offering many helpful comments that allowed me to make this thesis better. Thank you to Thomas and Magdalena for serving on my thesis committee, and for your energising words of encouragement during random encounters at the university and at conferences. Thank you to the rest of the LogiCS faculty for keeping this wonderful programme alive and thriving. Thank you to Clarissa and Doris for gently guiding me through the forest of graduation bureaucracy.

A very special and heartfelt thank-you goes to Anna, the founding programme manager of LogiCS. You moved mountains to secure my visa and provide me with a safe path to

Austria amid all the storms of 2022, and then did everything to welcome me into this country and this community. The regular coffee hours spent discussing our lives, the Christmas celebrations together, the seriousness you brought to taking pictures of us at every event—you did all this to make LogiCS into a family. The fact that you are no longer with us and cannot see us graduate is heartbreaking. But the fond memories of you, and of everything you did to help us grow, will be carried forward by many generations of LogiCS students.

Next, thank you to those who walk this path alongside me. First, to my PhD siblings, Domi and Josie. Domi, you were the first person I met in our office and the first Austrian I met in my life, and your aggressive friendliness made me feel at home right away. Thank you for serving as a constant cautionary tale of what overwork and excessive enthusiasm can do to a person, while somehow still managing to invent, organise, and often force me to join a million fun activities; for helping me shorten my bucket list by dyeing our hair together (and thank you to Tanja for carrying out the operation so masterfully!); for making me chuckle with your stupid stickers at every seminar we attend; and for creating the tradition of Christmas visits to Tulln, complete with the obligatory game of Overcooked, whose soundtrack accompanied the proofs of many theorems in this thesis. Josie, thank you for being a role model to me by showing that one can be a responsible, functional adult and a cool, fun person at the same time; for sharing your passions and showing your beautiful journal; for your emotional support and creating spaces where we could be open and vulnerable, both at cake meetings and on our special existential walks; for all our memorable activities, ranging from silent retreat meditations to loud techno clubbing. You both have defined my life in Austria, and I am incredibly happy that this PhD has given me such amazing friends, who I hope will remain in my life long after it is over.

My thanks also go to all the students and already-no-longer-students who have been part of our huge Theory and Logic group at various times: to Kees for introducing me to the fascinating world of jazz; to Blaž for boosting my confidence by overreacting to my jokes, and for his unconventional ideas about mathematical notation; to Robert for sharing stories from the real world over coffee; to Mirko for always sharing his wisdom and experience, and sometimes even his pasta; to Huimin for radiating energy, and for the coolest cosy socks and authentic sweets; to Matteo for bringing me to watch horror movies; to Giorgio and Dasha for hosting unforgettable bonkers parties; to Isabella and Jonatan for movie invitations and great movie discussions, and to Split for being such a great host; to Henri and Alessia for warm and welcoming board-game evenings; to Yoann for his irresistible enthusiasm for dragon boating; to Emery for sharing his enchanting fiction; to Thorsten for demonstrating us the possibilities of AI in research; to Glenn for his help and his patience with my complete technical cluelessness; to Stella for fun lessons in German politics and society; to Alexandra for book recommendations; to Xavier for determining which Vietnamese places are suitable for lunch and generously sharing his findings; and to Roman for his colossal enthusiasm for both logic and my various bizarre hairstyles. Thank you also to my fellow LogiCS students: especially to Davide for making

big conferences much less scary; to Tobias for organising and being the soul of regular after-work get-togethers; to Giovanni for boosting my confidence with his very generous classifications; to Liana for making me dream of Japan; and to Oliviero for teaching me some essential Gen Z memes. Thank you to Clemens and to my younger colleagues, Marit and Christian, for using your magical programming skills to turn my formulas into things that actually work.

I am immensely grateful to Emiliano, who so quickly turned my interest in his work into a fruitful collaboration—one that has been both intellectually stimulating and tremendous fun. Thank you for generously bringing me into your astonishingly diverse projects; for your lectures, which bring flavours of psychology and economics into logic; for your career advice; and for conversations over a burrito about marathon running and board games. Thank you for jumping through all the hoops of French bureaucracy to host me in Toulouse, and for devoting an incredible amount of time to me both there and afterwards. Scheduling a call on your birthday to discuss counterfactuals and belief revision was almost too much devotion! Thank you also to all the amazing people at IRIT who made me feel so welcome in Toulouse—especially to Carlos for revealing whole new dimensions of Hanabi, to Hans for helping me find the right path, and to Andreas for his kindness and wonderfully laid-back spirit.

elFinally, a special thank-you to my dearest friends far away. To Sanya and Kris, for sharing your passion for music and always finding the most amazing ways to reboot my brain whenever it gets stuck on a problem. I owe you the ideas behind Chapter 7 of this thesis, which emerged this way after three days of partying together at Beonix. To Dimka, Marina, and Arty, for giving me a warm and peaceful place in lovely Münster to relax and recharge. To Masha and Kravch, for untangling the knots of my recurring existential crises during our hours-long calls. Last but not least, to my mom, for always being there for me and for meeting both my ramblings about the intricacies of yet another semantic completeness proof and my stories of everyday life in Austria with the same unfailing enthusiasm.

Kurzfassung

Konditionale Aussagen – Aussagen der Form „wenn A, dann B“ – treten in einer bemerkenswerten Vielfalt von Kontexten auf: In der Philosophie werden kontrafaktische Konditionale zur Analyse von Kausalität und wissenschaftlicher Erklärung verwendet; in Recht und Ethik drücken bedingte Normen kontextabhängige Verpflichtungen und Erlaubnisse aus; in der Künstlichen Intelligenz bilden Konditionale die formale Grundlage für defeasible und nichtmonotone Schlussverfahren über typische Situationen und unvollständige Informationen. Diese Vielseitigkeit hat zu einer reichen Landschaft logischer Formalismen geführt, die jeweils auf eine bestimmte Interpretation von Konditionalen und bestimmte Anwendungsszenarien zugeschnitten sind. Obwohl viele dieser Formalismen innerhalb ihrer jeweiligen Forschungstraditionen eingehend untersucht wurden, sind ihre Eigenschaften und wechselseitigen Beziehungen bislang nur teilweise verstanden, und zahlreiche offene Fragen bestehen fort. Diese Dissertation untersucht konditionale Logiken aus der Perspektive analytischer Methoden. Solche Methoden erfassen logische Folgebeziehungen durch syntaktische oder semantische Konstruktionen, deren Struktur eng durch die zu untersuchenden Formeln bestimmt wird, wodurch sie sich leichter konstruieren und analysieren lassen. Es werden neue analytische Charakterisierungen sowohl in beweistheoretischer als auch in modelltheoretischer Hinsicht für mehrere bedeutende konditionale Formalismen entwickelt, die drei zentrale Interpretationen von Konditionalen abdecken: kontrafaktische Aussagen, Verpflichtungen sowie defeasible Aussagen über typische Situationen. Diese Charakterisierungen liefern einheitliche und modulare Darstellungen der betrachteten Logikklassen und decken neue Verbindungen zwischen verschiedenen konditionalen Formalismen und etablierten Modallogiken auf. Darüber hinaus schaffen sie eine solide rechnerische Grundlage, die es ermöglicht, die exakte Komplexität der betrachteten Logiken und ihrer Sprachfragmente zu bestimmen, und dienen zugleich als Basis für die Entwicklung effizienter Werkzeuge zum automatisierten Schließen. Die Ergebnisse zeigen, dass analytische Methoden eine fruchtbare Perspektive auf das konditionale Schließen eröffnen, die sowohl zur Lösung offener Probleme als auch zum Transfer von Techniken und Resultaten zwischen unterschiedlichen formalen Rahmenwerken beiträgt.

Abstract

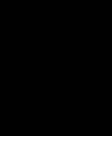
Conditionals — statements of the form “if A then B ” — arise across a remarkably diverse range of contexts: in philosophy, counterfactual conditionals are used in the analysis of causation and scientific explanation; in law and ethics, conditional norms express context-dependent obligations and permissions; in artificial intelligence, conditionals provide the formal basis for defeasible and non-monotonic reasoning about typical situations and incomplete information. This versatility has given rise to a rich landscape of logical formalisms, each tailored to a specific interpretation of conditionals and application scenarios. While many of these formalisms have been studied extensively within their respective research traditions, their properties and mutual relations remain unevenly understood, with a number of open questions persisting across the field. This thesis studies conditional logics through the lens of analytic methods, which capture logical entailment via syntactic or semantic constructions whose structure is tightly controlled by the formulas under evaluation, making it easier to construct and analyse them. We develop new analytic characterizations in both proof-theoretic and model-theoretic terms for several prominent conditional formalisms spanning three major interpretations of conditionals: counterfactual statements, obligations, and defeasible statements about typical situations. These characterizations provide uniform and modular representations of the considered classes of logics and uncover new connections between different conditional formalisms and well-established modal logics. They also provide a solid computational foundation, enabling us to determine the exact computational complexity of the considered logics and their language fragments, and serve as the basis for developing efficient automated reasoning tools. This demonstrates that analytic methods offer a fruitful perspective on conditional reasoning, allowing to tackle open problems and transfer techniques and results across different settings.

Contents

Kurzfassung	ix
Abstract	xi
Contents	xiii
1 Introduction	1
1.1 Motivation and goals of the thesis	1
1.2 Analytic methods in logic	4
1.3 Overview of results and covered formalisms	6
1.4 Publications	9
1.5 Structure of the thesis	11
2 Semantical and Proof-Theoretic Foundations	15
2.1 Classical and normal modal logics	15
2.2 Sequent calculi for classical and normal modal logics	18
2.3 Hypersequent calculi	20
2.4 Antisequent calculus	25
I Counterfactuals	27
3 Preferential Conditional Logic	29
3.1 Background	30
3.2 State-of-the-art	30
3.3 Preliminaries: Preferential models and PCL	31
3.4 GL -based sequent calculus for PCL	33
3.5 Bridge : Embedding of PCL into GL+K	40
II Norms	43
4 Preference-Based Deontic Logics	45
4.1 Background	46
	xiii

4.2	State of the art	47
4.3	Formal preliminaries: Åqvist's logics	48
4.4	LEGO model constructions for Åqvist's logics	51
4.5	Alternative frame conditions for Åqvist's logics	67
4.6	Complexity: Åqvist's logics are co-NP-complete	69
4.7	Prover: Deo-SMT	70
4.8	GL-based hypersequent calculi for Åqvist's logics	70
4.9	Bridge: Embeddings of Åqvist's logics into normal bimodal logics . .	81
5	Norm-Based Deontic Logics	85
5.1	Background	86
5.2	State of the art	87
5.3	Preliminaries: Axioms of the core I/O logics	88
5.4	Sequent calculi and separation principles for I/O logics	90
5.5	Bridge: Reducing deontic I/O entailment to causal I/O entailment .	112
5.6	Semantics of Input/Output models	112
5.7	Bridge: I/O models and existing semantical descriptions for I/O logics	115
5.8	Bridge: Embeddings of I/O logics into classical and normal modal logics	116
5.9	Bridge: Reducing entailment in Åqvist's logic E to I/O entailment .	121
5.10	Prover: IOentailment	123
5.11	Complexity: I/O logics and language fragments	124
5.12	Complexity: Bochman's actual cause is in NP	128
III	Typicality	131
6	Locally non-monotonic reasoning	133
6.1	Background	134
6.2	State of the art	135
6.3	Preliminaries: KLM hierarchy	136
6.4	Analytic version of smoothness condition	138
6.5	Saturating sequent calculi for KLM logics	140
6.6	Complexity: all KLM logics are co-NP-complete	150
6.7	Prover: KLMPpropagator	152
7	Globally non-monotonic reasoning	155
7.1	Background	156
7.2	State of the art	156
7.3	Preliminaries: Two perspectives on rational closure	157
7.4	Hypersequent calculus for saturated reasoning in ranked models	160
7.5	Hypersequent calculus for rational closure	166
8	Conclusion	173
8.1	Related conditional formalisms	173

8.2 Future perspectives	176
Overview of Generative AI Tools Used	179
List of Figures	181
List of Tables	183
List of Algorithms	185
Bibliography	187



Introduction

1.1 Motivation and goals of the thesis

This thesis is devoted to advancing theory and tools of logical formalisms for reasoning about conditionals, namely statements of the form “if A then B ”. Such statements arise in many forms of conditional reasoning and usually cannot be adequately captured by the material implication of classical logic. This observation has played a significant role in the development of modern logic: early attempts to capture stronger notions of implication [LL32] gave rise to the central modal logics **S4** and **S5**. Subsequently, the study of conditionals has become a meeting point for several areas of contemporary logic, ranging from philosophical logic to knowledge representation. In analytic philosophy, counterfactual conditionals (“if it were the case that A , then B would hold”) have been extensively studied and have had a profound influence on the analysis of causation [Lew86], explanation [Woo03], and scientific laws [Lew79]. Conditionals also plays a central role for normative reasoning, as norms (obligations, prohibitions and permissions) are usually contextual (“under the condition A , B is obligatory”) and are formalized using conditionals. They are essential for reasoning about contrary-to-duty obligations (obligations that come into force when other obligations are violated) and for avoiding the deontic paradoxes [GHP⁺13] that beset Standard Deontic Logic [vW51], which represents obligation as a unary modality. In artificial intelligence and knowledge representation, conditionals provide a natural formal basis for reasoning with typical situations (“ A typically implies B ”) and exceptions –often referred to as defeasible or non-monotonic reasoning– and for representing knowledge under incomplete information in intelligent systems [KI19].

As a result of these diverse motivations and applications, a rich landscape of conditional formalisms has emerged, each tailored to particular interpretations of conditionals and specific domains of application.

Since conditionals play important roles across several fields, the corresponding formalisms have often developed within different research communities and methodological traditions. As a consequence, the theory of conditional logics remains fragmented despite the maturity of the field. Theoretical results are often scattered and tailored to particular systems. Methods developed for one logic do not always extend smoothly even to closely related variants within the same family, and they rarely transfer across different interpretations of conditionals, leaving structural connections between the logics only partially understood. These difficulties are particularly visible in the study of computational aspects, such as the complexity of entailment and algorithmic decision procedures. These aspects are crucial for the automation and application of conditional reasoning for practical tasks; yet for many conditional logics, they remain only partially understood, and the available procedures are usually far from the theoretical optimum.

To address these issues, this thesis studies conditional logics through the lens of *analytic characterizations*. Roughly speaking, analytic characterizations aim to represent entailment in a given logic using constructions that are tightly controlled by the formulas occurring in the entailment. In proof-theoretic settings, this idea is typically captured by proof calculi satisfying the subformula property, ensuring that derivations are built only from components appearing in the formulas of the derived statements. The same analytic perspective can also be brought to the semantic side through countermodel constructions, where bounded models witnessing non-entailment are systematically built from the formulas occurring in the statement. By making the structure of entailment explicit and restricted in this way, analytic characterizations provide a useful framework for comparing different conditional formalisms, analyzing their computational behavior, and developing automated decision procedures.

Analytic methods have already been applied to a number of conditional logics in the literature. However, these results are often scattered and typically tailored to particular systems, sometimes covering only fragments or individual members of a given logical family. In this thesis, we develop new methods to obtain analytic characterizations for conditional logics guided by several desiderata. First, the characterizations should capture all the standard systems within a given logical family using the same general construction (uniformity), while allowing individual systems to be obtained by adding or modifying the specific components—such as inference rules in a calculus or semantic building blocks in a model construction—corresponding to their distinguishing properties (modularity). In this way, closely related logics can be studied within a common framework rather than by means of separate ad hoc constructions. Second, the characterizations should support a close interaction between proof-theoretic and semantic perspectives, providing a unified understanding of the logic and allowing one to systematically obtain witnesses for both entailment, in the form of proof-theoretic derivations, and non-entailment, in the form of semantic countermodels. Last but not least, the characterizations should support the following applications:

- **Complexity:** providing tools for determining the exact complexity of the entailment

problem in the logic (usually referred to simply as the complexity of the logic) and for identifying tractable fragments of the language.

- **Provers:** facilitating automated reasoning using “reductionist approach”, in which existence of an entailment witness (e.g. derivation or countermodel) is encoded as an instance of some standard problems such as SAT, enabling lightweight and efficient implementations that exploit powerful general-purpose solvers.
- **Bridges:** uncovering new connections between different conditional formalisms and with other logical systems, usually in the form of embeddings, i.e., translations preserving entailment in both directions.

We do not aim to cover as many conditional formalisms as possible here. Instead, we consider three common interpretations of conditionals: counterfactual reasoning, normative reasoning, and reasoning about typicality; we focus on five well-known formalisms (a logic or a family of logics), each being a representative of one of the standard approaches within these areas. These five formalisms are:

1. Preferential Conditional Logic [Bur81] representing the preferential semantics tradition for counterfactual reasoning;
2. Åqvist’s family of logics [Åqv02] for preference-based normative reasoning;
3. Input/Output family of logics [MvdT00] for rule-based normative reasoning;
4. the KLM family of logics [KLM90] for locally non-monotonic¹ reasoning about typicality;
5. the rational closure [LM92] as a canonical globally non-monotonic conditional entailment operator.

These formalisms are extensively discussed in the respective chapters. Each of these formalisms embodies distinctive semantic and methodological ideas, and the analytic characterizations developed in this thesis are tailored to their specific features. At the same time, all our characterizations follow a common methodological perspective, moving back and forth between proof-theoretic and semantic viewpoints and transferring ideas between the different settings. This allows us to highlight both the diversity of conditional formalisms and the methodological connections between them.

¹In non-monotonic reasoning, two levels of non-monotonicity are distinguished [Boc06]: local, where entailment may be violated under strengthening the antecedent, and global, where entailment may be violated under extension of the knowledge base.

1.2 Analytic methods in logic

In this thesis, we use the term *analytic* in a broad sense for logical methods in which the reasoning process is guided and constrained by input formulas, rather than involving arbitrary formulas unrelated to the input. The idea of analytic reasoning in logic goes back to Gottfried Wilhelm Leibniz, who distinguished truths of reason from truths of fact. He viewed truths of reason as analytic, since their validity can be uncovered by analysing the involved concepts, rather than by appeal to external facts, an idea that later inspired his vision of a calculus ratiocinator, a symbolic system for mechanical reasoning.

In modern logic, these ideas are formally realized in the area of structural proof theory, with the introduction of the notion of analytic calculi. A central framework in this area is Gentzen's sequent calculus [Gen34], in which proofs are represented as derivations manipulating sequents, i.e., collections of assumptions and conclusions. Sequent-style calculi make it possible to achieve analyticity in its strongest form, the *subformula property*, ensuring that derivations involve only subformulas of the statements being proved. More generally, we will also consider a proof calculus analytic when it satisfies a relaxed subformula property, allowing a finite collection of auxiliary formulas systematically constructed from these subformulas. By contrast, standard Hilbert-style proof calculi, based on unrestricted applications of modus ponens, are non-analytic in this sense, since intermediate formulas in a derivation need not be determined by the formula being proved. The subformula property is often achieved by proving the admissibility or elimination of the cut rule, which simulates modus ponens in Hilbert calculi and is therefore crucial for establishing the completeness of sequent calculi with respect to the underlying logic. The restricted form of cut-free derivations, namely derivations without the cut rule, makes sequent-style calculi a particularly convenient framework for analyzing the behaviour of entailment in the logic, and to prove meta-logical results (decidability, complexity, etc.). For this reason, the proof-theoretic perspective on conditional logics adopted in this thesis will be based on sequent-style calculi.

Originally introduced for classical first-order logic, sequent calculi were later extended to a wide range of (propositional) modal logics too, see, e.g. [Wan98, Pog11]. However, the expressive power of ordinary sequents is limited in this setting. In particular, for the central modal logic **S5**, no cut-free sequent calculus is known. To overcome these limitations, several generalizations of the sequent framework have been developed. The simplest of these are hypersequents [Avr96b], where derivations manipulate disjunctive lists of sequents rather than single sequents. Hypersequents already suffice to obtain analytic calculi for a number of modal logics, including **S5**. More expressive frameworks include nested sequents [Brü09], display calculi [Bel82], and labeled sequents [Gab96, Neg05], which incorporate richer structural or semantic information into derivations. While these formalisms increase expressive power, they also make derivations more complex, and make it more difficult both to establish meta-logical results for the formalized logics and to extract effective methods for automated deduction [LCG⁺23]. For this reason, the proof-theoretic developments in this thesis will focus on the simpler sequent-style frameworks of sequents and hypersequents.

Analytic (or compositional) methods are not confined to proof theory: a rich tradition of such approaches also exists on the semantic side. In this setting, analyticity means that the building blocks of a model and its relevant structure are determined by a finite closure of the formulas under consideration. Thus, the structure of a formula guides the construction of a model sufficient to establish satisfiability or provide a counterexample, rather than requiring an unrestricted construction, as in standard canonical-model constructions, whose worlds are typically defined as maximally consistent sets over the entire language. Early examples of analytic semantic methods include Beth’s refutation procedures [Bet55], which decompose formulas to construct countermodels witnessing non-derivability, and Hintikka’s theory of model sets [Hin62], based on closure conditions for sets of subformulas from which models can be generated. Kripke adapted this perspective to modal logic in his completeness proofs [Kri66], using the structure of formulas to inductively construct Kripke frames. Analytic model constructions later became central in the study of computational properties of modal logics. In particular, Ladner’s landmark complexity results [Lad77] rely on a memory-efficient recursive model construction based on subformula analysis. More general versions of this approach were developed by Friedman and Halpern [FH94], whose small-model constructions for conditional logics provide a modular method for synthesizing bounded models from components inductively obtained for subformulas.

Although proof-theoretic and semantic characterizations are often presented as distinct or even opposing perspectives, in the analytic setting, they become closely intertwined. Specifically, proof search in analytic calculi can be viewed semantically as a systematic search for countermodels, where failed proof attempts give rise to countermodels reconstructed from the structure of the derivation [NvP11]. This perspective is particularly natural in “saturating” sequent and hypersequent calculi (also called Kleene-style calculi [KL18]), where derivations proceed by monotonically accumulating information until a saturated state is reached, from which a countermodel can be directly extracted. A similar view underlies the formalism of analytic tableaux [Smu68], which provide a proof-theoretic formulation of Beth’s refutation procedures. In this way, proof search and countermodel construction become part of a single systematic process. Moreover, this semantic reading of sequent calculi provides elegant semantic methods for establishing completeness of a calculus, without relying on involved syntactic transformations or axiomatic systems.

In this thesis, we will adopt such a two-sided analytic approach to the study of conditional logics, developing both proof-theoretic and semantic characterizations that inform and refine one another. By allowing the development on different sides to guide each other, we ensure that the resulting formalisms are naturally aligned and provide a holistic understanding of the logic, its structural and computational properties.

1.3 Overview of results and covered formalisms

This thesis investigates the following known formalisms, classified below according to interpretation of the conditionals:

Counterfactual: Preferential Conditional Logic (**PCL**) [Bur81],

Normative:

- Preference-based: Åqvist’s family of deontic logics [Åqv02], comprising the systems **E**, **F**, **F_{CM}**, and **G**;
- Norm-based: the core deontic and causal Input/Output (I/O) logics, comprising the original systems **OUT₁–OUT₄** [MvdT00], the causal systems **OUT₁[⊥]–OUT₄[⊥]** [Boc04], and the throughput systems **OUT₁⁺–OUT₄⁺** [MvdT01];

Typicality

- Locally non-monotonic: the KLM hierarchy [KLM90] of non-monotonic logics, comprising the systems **C**, **CL**, **P**, **CM**, and **M**;
- Globally non-monotonic: the rational closure operator **RC** of conditional entailment from conditional knowledge bases [LM92].

For each of these formalisms, we develop novel analytic characterizations combining proof-theoretic and semantic perspectives, and use them to study the formalisms, their complexity, connections with other logical systems, and to develop automated deduction methods. A visual overview of the covered formalisms and obtained results is given in Fig. 1.1. The main contributions of the thesis can be grouped into several categories, which we summarize below.

Analytic Calculi. On the proof-theoretic side, the characterizations introduced in this thesis take the form of sequent calculi (for **PCL**, I/O logics, and KLM logics) and of hypersequent calculi (for Åqvist’s logics and rational closure). Relative to existing proof systems, the contributions range from the first analytic characterizations, where none were previously available (for rational closure and the Input/Output systems **OUT₂** and **OUT₄**), through substantial simplifications of existing calculi (for **PCL** and Åqvist’s system **F**), to alternative uniform proof-theoretic frameworks that avoid system-specific rules or auxiliary proof-search machinery present of earlier approaches (for the Åqvist family and the KLM logics). Each family of logics is developed in a modular way, combining a common rule base with additional rules specific to the individual systems; additionally, the calculus for rational closure also modularly covers the KLM logic **P** to uniformly represent both local and global dimensions of non-monotonic reasoning. All calculi admit a semantic interpretation, and countermodels can be extracted from failed derivations. Proof search in the calculi is shown to match the optimal computational complexity of the corresponding logics, with the exception of Åqvist’s systems, where computational optimality is instead achieved using semantic methods.

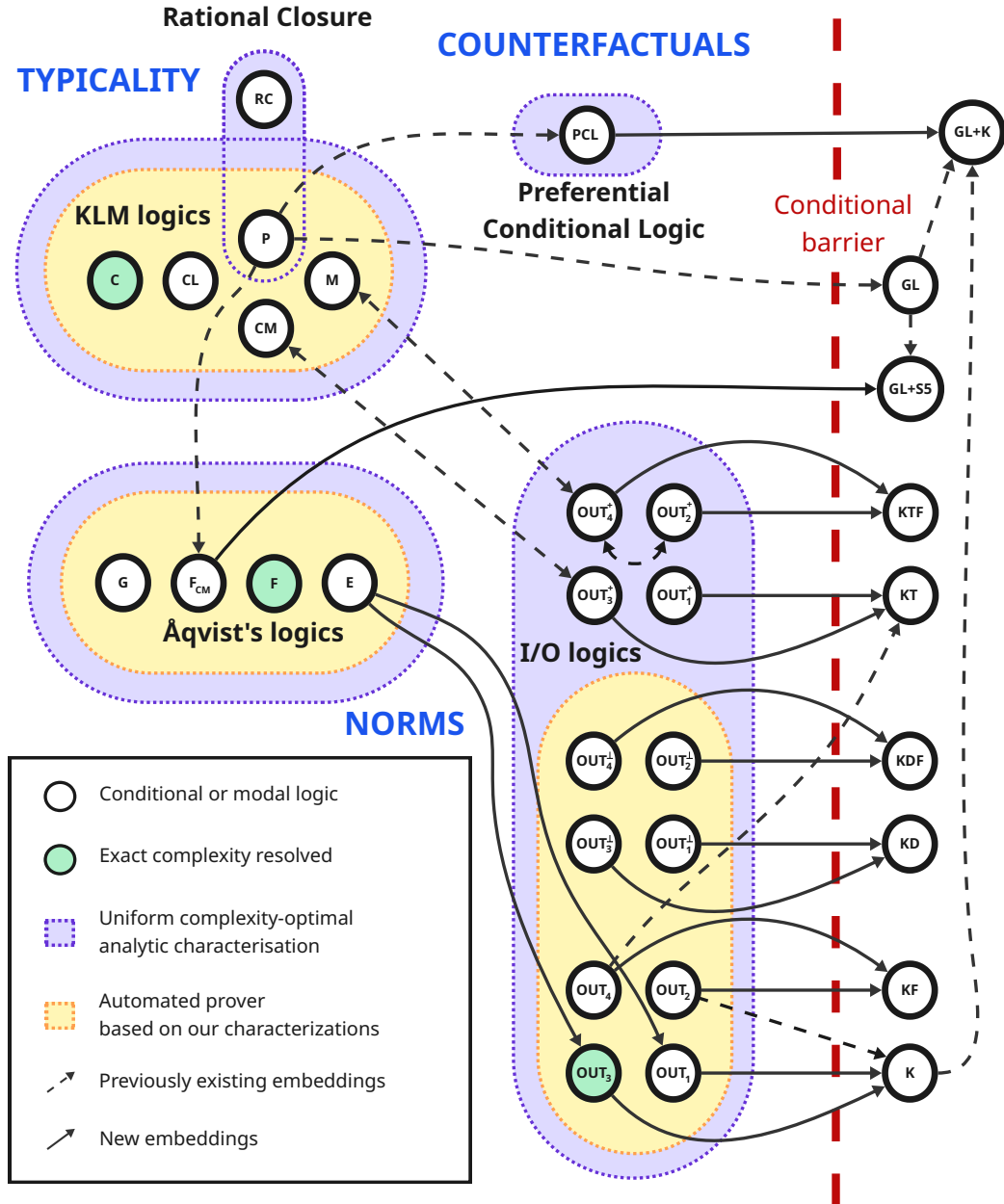


Figure 1.1: Map of the covered logics and obtained results

Analytic Semantics. On the semantic side, we provide new analytic semantic characterizations for Åqvist’s logics (via small-model constructions and alternative frame conditions), Input/Output logics (via a new kind of possible-world semantics), and the KLM logic **C** (via an alternative frame condition). These characterizations either serve as the basis for the development of the corresponding proof calculi (for Åqvist’s and KLM logics) or are, conversely, derived from the proof-theoretic analysis itself (for I/O logics). In all cases, the calculi provide a systematic and constructive method for extracting countermodels with respect to the corresponding semantics from failed derivations. In particular, for **PCL** this yields a systematic way to reconstruct the Friedman–Halpern small-models constructions [FH94], whose existence had previously only been established non-constructively.

Computational Complexity. The computational complexity of all considered formalisms has been previously studied, yet the picture remained incomplete for several of them. While exact complexity classes were established for **PCL** and rational closure (PSPACE [FH94] and P^{NP} [EL00] respectively), the complexity of Åqvist’s logics, I/O logics, and KLM logics was only partially understood, with certain individual members of those families (Åqvist’s logic **F**, I/O logic **OUT₃**, and KLM logic **C**) resisting precise classification by the suggested methods. Our analytic characterizations provide a natural and uniform way to study complexity, yielding a complete picture of complexity classes for all considered logics, including the complexity of important language fragments of the I/O and KLM families.

Automated deduction. Our characterizations also admit natural encodings into standard computational problems — in most cases, propositional satisfiability — enabling lightweight and efficient automated deduction by leveraging powerful general-purpose solvers. Such tools based on our characterizations were implemented by collaborators for Åqvist’s logics (**DeoSMT** by Christian Köll [Köl25]), for I/O logics (**IOentailment** by Marit Einböck [Ein24]), and for KLM logics (**KLMPpropagator** by Clemens Eisenhofer [CER24]). Initial evaluations suggest that these tools outperform existing provers based on direct proof search or on encodings into computationally heavier formalisms, such as Higher-Order Logic.

Bridges. The considered conditional formalisms studied in this thesis were largely developed and studied independently. There are, nevertheless, some known intersections between them: the KLM logic **P** coincides with the flat fragment (i.e. without nesting of conditionals) of **PCL** and of Åqvist’s logics **F_{CM}** and **G**, the I/O logic **OUT₃⁺** coincides with the KLM logic **CM**, and classical logic also appears both in I/O family (where both **OUT₂⁺** and **OUT₄⁺** collapse to it) and in the KLM hierarchy (under the name **M** for “monotonic reasoning”). There are also several isolated results connecting these formalisms with normal modal logics: the I/O logics **OUT₂** and **OUT₄** were embedded into **K** and **T** [MvdT00], respectively, with similar embeddings later obtained for their causal counterparts **OUT₂[⊥]** and **OUT₄[⊥]** [Boc04]; conditionals in Åqvist’s logic **E** can be

represented using combinations of modalities from **K** and **S5** [COP22], while conditionals in the KLM logic **P** admit representations via modalities of **K** and the Gödel Löb logic **GL** [GGOP09]. Beyond such isolated correspondences, however, relatively little attention has been devoted to systematic translations and structural connections between these formalisms. Our analytic characterizations provide an abstract and modular description of the considered logics, making such connections easier to identify and study. They allow us both to re-establish known bridges described above and to uncover new ones. In particular, we simplify the known connection between the KLM logic **P** and the provability logic **GL**, and extend it to **PCL** and Åqvist’s logics. For the I/O family, we obtain new uniform embeddings into classical and normal modal logics that cover all core systems, subsuming previously known embeddings that applied only to parts of the family. Finally, we establish genuinely new connections between different conditional formalisms, developing simple syntactic and semantic links between the original (deontic) and causal I/O logics and showing that entailment in (preference-based) Åqvist’s logic **E** can be represented in the (norm-based) I/O logic **OUT₃[⊥]**.

Solution to open problems. Our characterizations provide a uniform and comprehensive picture of the considered formalisms, encompassing many scattered results and offering new findings, including solutions to the following open problems from the literature:

- proving that Åqvist’s system **F** is characterized by finite and acyclic preference models, as conjectured in [COP⁺23];
- providing a semantical link between procedural and bimodal characterizations of I/O logics, declared missing in [vdTP13];
- developing analytic sequent calculi for I/O logics **OUT₂** and **OUT₄**, which resisted the previous approach suggested in [Lel21];
- constructing embeddings of I/O logics **OUT₁** and **OUT₃** into normal modal logics, conjectured not possible in [MvdT00];
- establishing co-NP-completeness of Åqvist’s system **F**, conjectured in [COP⁺23] (NEXPTIME complexity upper bound is provided there);
- establishing co-NP-completeness of **OUT₃**, whose complexity was previously estimated to be between co-NP and P^{NP} in [SR17];
- establishing co-NP-completeness of KLM logic **C**, whose complexity was previously estimated with the upper bounds NEXPTIME [GGOP09] and PSPACE [AOP16].

1.4 Publications

The main technical chapters of this thesis are based on the following publications, extended with original contributions and additional connections, with the aim of providing a uniform and comprehensive account of the analytic methods developed for conditional logics.

1. Agata Ciabattoni, and Dmitry Rozplokhas. *Streamlining Input/Output Logics with Sequent Calculi*. 20th International Conference on Principles of Knowledge Representation and Reasoning (KR), 2023.
2. Dmitry Rozplokhas. *LEGO-Like Small Model Constructions for Åqvist's Logics*. 15th International Conference on Advances in Modal Logic (AiML), 2024.
3. Agata Ciabattoni, Clemens Eisenhofer, and Dmitry Rozplokhas. *Strongly Analytic Calculi for KLM Logics with SMT-Based Prover*. 21st International Conference on Principles of Knowledge Representation and Reasoning (KR), 2024.
4. Agata Ciabattoni, Dmitry Rozplokhas, and Matteo Tesi. *GL-Based Calculi for PCL and Its Deontic Cousin*. 19th European Conference on Logics in Artificial Intelligence (JELIA), 2025.
5. Christian Köll, Agata Ciabattoni, and Dmitry Rozplokhas. *SMT-Based Deontic Reasoning for Åqvist Logics*. 35th International Joint Conference on Automated Reasoning (IJCAR), 2026, to appear.
6. Agata Ciabattoni, and Dmitry Rozplokhas. *Streamlining Input/Output Logics: Sequent Calculi, Semantics, and Computational Complexity*. Submitted.²
7. Dmitry Rozplokhas. *Analytic Proof Theory for Rational Closure*. Submitted.

Additionally, the following publications produced during the PhD are related to the results presented in this thesis and are discussed in the relevant sections.

8. Agata Ciabattoni, Nicola Olivetti, Xavier Parent, Revantha Ramanayake, and Dmitry Rozplokhas. *Analytic Proof Theory for Åqvist's System F*. 16th International Conference on Deontic Logic and Normative Systems (DEON), 2023.
9. Emiliano Lorini, and Dmitry Rozplokhas. *Graded Distributed Belief*. 21st Conference on Theoretical Aspects of Rationality and Knowledge (TARK), 2025.
10. Carlos Aguilera-Ventura, Xinghan Liu, Emiliano Lorini, and Dmitry Rozplokhas. *A Non-Interventionist Approach to Causal Reasoning based on Lewisian Counterfactuals*. 34th International Joint Conference on Artificial Intelligence (IJCAI), 2025.

²Invited journal extension of the KR'23 paper.

11. Agata Ciabattini, Josephine Dik, Emiliano Lorini, Dominik Pichler, and Dmitry Rozplokhas. *From Explicit Allowances to Defeasible Deontic Operators: A Modal View*. 26th International Conference on Principles and Practice of Multi-Agent Systems (PRIMA), 2025.
12. Emiliano Lorini, and Dmitry Rozplokhas. *Merging Beliefs and Counterfactuals*. 35th International Joint Conference on Artificial Intelligence (IJCAI), 2026, to appear.

1.5 Structure of the thesis

The thesis consists of five main technical chapters, each devoted to a specific conditional formalism; in addition there is an introductory chapter, a chapter of logical and proof-theoretic preliminaries, and a concluding chapter. The technical chapters are grouped into three parts, each devoted to one interpretation of conditionals:

- counterfactual reasoning,
- normative reasoning,
- typicality-based reasoning.

Each chapter follows a common structure. It begins with a background section, outlining the historical context of the given approach to conditional reasoning and introducing the formalism under consideration, and a “State of the art” section, summarizing the relevant existing results for this formalism, followed by a section introducing the necessary technical preliminaries. The core of each chapter is then devoted to developing new analytic characterizations of the formalism in question, with particular emphasis on the interaction between proof-theoretic and semantic perspectives, and to applying these characterizations to the study of the formalism. As part of this study, sections corresponding to the three main application directions are explicitly marked in their titles: “**Complexity**” sections establish exact computational complexity results and complete previously incomplete classifications; “**Prover**” sections describe automated reasoning tools implemented on the basis of the proposed characterizations by collaborators; and “**Bridge**” sections present new connections between the studied formalism and other logical systems, typically in the form of embeddings or syntactic “bridging conditions”.

Chapter 2 collects the formal preliminaries required for the subsequent developments, introducing normal modal logics together with the sequent, hypersequent, and antisequent frameworks that provide the basis for the analytic methods developed throughout the thesis.

Chapter 3 investigates the tradition of counterfactual reasoning based on preferential models, represented by Preferential Conditional Logic (**PCL**) of Burgess. We develop a sequent calculus for this logic by combining two key insights: extending the connection

to Gödel-Löb provability logic **GL** observed in [GGOP09] for the KLM logic **P** (corresponding to the unnested fragment of **PCL**), and the small model constructions of Friedman and Halpern [FH94]. The calculus achieves optimal computational complexity of proof search and provides a constructive procedure for obtaining Friedman–Halpern small models. We also use the calculus to bridge (nested) conditional logics and **GL**.

The technical content of this chapter is based on paper #4.

Chapter 4 investigates conditional deontic reasoning in the preferential approach of Åqvist’s systems [Åqv02] with increasing deductive power: namely **E**, **F**, **F_{CM}**, **G**. All these logics contain **S5** as a fragment, making cut-free characterizations via ordinary sequent calculi inapplicable. Since the weaker systems **E** and **F** lack key model-theoretic properties used in the analysis of **PCL**—transitivity and smoothness—the Friedman–Halpern construction [FH94] does not apply to them either. We therefore introduce a new analytic model construction method, called LEGO-models, based on composing elementary semantic building blocks, which applies uniformly across the entire Åqvist family. Using this framework, we obtain new semantic characterizations of **F**, establish co-NP-completeness for all four Åqvist systems (resolving the exact complexity of **F**), and develop a modular hypersequent framework for **F**, **F_{CM}**, and **G** extending the existing calculus for **E** [COP22]. As in the case of **PCL**, the resulting calculi also yield embeddings into normal modal logics.

The technical content of this chapter is based on papers #2, #4, #5, and includes description of paper #8. Sec. 4.8.2 and Sec. 4.8.3, presenting analytic calculi for **G** and **F**, constitute original unpublished work.

Chapter 5 turns to the norm-based tradition of deontic reasoning, represented by the framework of Input/Output (I/O) logics. We study twelve standard I/O systems, including the original logics, their throughput variants, and modifications for causal reasoning. We develop modular analytic sequent calculi for all of them, substantially extending the limited range of previously known proof systems. Central to this development are new separation principles characterizing I/O entailment via families of classical entailments over partitions of the norm base. These principles yield a simple bridge condition connecting the original and causal variants and, when read contrapositively, give rise to a new possible-world semantics uniformly covering all twelve systems. This semantics generalizes earlier semantic approaches, clarifying the relationship between the different variants of I/O logic, and enabling new connections with Åqvist’s preference-based deontic logics. Building on our analytic characterizations, we establish embeddings into classical logic and embeddings into modal logics, extending previous results of Makinson and Bochman, encode conditional entailment in Åqvist’s logic **E** into **OUT₁** and **OUT₃**, and obtain a complete classification of the computational complexity of the considered systems (resolving the exact complexity of **OUT₃[⊥]**) and important language fragments.

The technical content of this chapter is based on paper #1 and its journal extension #6. Section 5.9, presenting an encoding of conditional entailment in Åqvist’s logic **E** into I/O logics, constitutes original, unpublished work.

Chapter 6 studies the KLM framework for non-monotonic reasoning, viewed as a family of flat (i.e. without nesting of conditionals) conditional logics capturing local non-monotonicity. First, we suggest a modified semantics for the weakest KLM logic **C**, based on “analytic” relaxed version of the smoothness property for preference models. Building on this approach, we generalize the separation principles of Chapter 5 to separation graphs and develop “strongly analytic” sequent calculi based on saturation processes over conditional dependency graphs, thereby obtaining a uniform proof-theoretic perspective on the KLM family. As in the case of I/O logics, the resulting framework yields a complete classification of complexity classes for all systems (resolving the exact complexity of **C**) and important language fragments.

The technical content of this chapter is based on paper #3.

Chapter 7 is devoted to the analytic characterization of rational closure, the canonical approach to globally non-monotonic entailment from conditional knowledge bases. While it has been extensively studied from semantic and constructive perspectives, no analytic calculus was previously known. Having such a calculus is important, since derivations can provide formal explanations for entailment and be used to study structural properties of rational closure, such as interpolation. To address this, we combine proof-theoretic techniques developed in earlier chapters — in particular, the **GL**-based hypersequent methods of Chapter 4 — together with other established proof-theoretic techniques, such as analytic cuts and antisequents. This yields a modular hypersequent calculus covering both the locally non-monotonic logic **P** and globally non-monotonic rational closure, with the transition between them governed by a single inference rule. In the course of establishing completeness, we obtain a direct correspondence between the proof-theoretic, semantic, and constructive characterizations of rational closure.

The technical content of this chapter is based on paper #7 under submission.

In the concluding Chapter 8, we discuss several related conditional formalisms and the applicability of the analytic methods introduced in the thesis to them, and outline possible directions for the further development and application of these methods.

This chapter includes descriptions of the papers #9, #10, #11, #12.

Note on reading order. Chapter 2 provides the common formal background for the technical chapters; whenever material from this chapter is used later, the relevant section or result is explicitly referenced. The main technical chapters are mostly self-contained, except for sections establishing cross-chapter bridges, where any necessary background from earlier chapters is likewise referenced explicitly. Nevertheless, reading the chapters sequentially may help to appreciate the methodological development of the thesis. In particular, two methodological threads recur across several chapters: the **GL**-based proof-theoretic approach developed for **PCL** (Chapter 3) is adapted to Åqvist’s logics (Chapter 4) and later reappears in the treatment of rational closure (Chapter 7), while the separation principles introduced for I/O logics (Chapter 5) are developed further into separation graphs for the KLM family (Chapter 6). The latter chapters also lead

naturally from locally non-monotonic reasoning in Chapter 6 to globally non-monotonic reasoning in Chapter 7. Thus, the proposed order helps build intuition for the connections and methodological development across the thesis, but is not required for understanding the results in individual chapters.

Semantical and Proof-Theoretic Foundations

This chapter collects the formal preliminaries on modal logic and analytic proof-theory used throughout the thesis. We begin in Sec. 2.1 by introducing the relational semantics for normal modal logics together with several standard modal systems relevant to our later developments for conditional logics. Next, Sec. 2.2 presents Gentzen's sequent framework for classical logic and its extensions to normal modal logics. In Sec. 2.3, we introduce hypersequents as a simplest generalization of sequents and discuss the Kleene style of derivations using a hypersequent calculus for **S5** as a representative. Finally, Sec. 2.4 presents the antisequent calculus, a proof-theoretic framework for deriving non-validity in classical logic that will later serve as a building block in our treatment of globally non-monotonic reasoning.

2.1 Classical and normal modal logics

We begin by fixing some basic terminology. A (propositional) *logical language* $\mathcal{L}$ is a set of formulas generated from a given collection of propositional atoms, logical connectives and modalities. In this thesis, we use capital Latin letters to denote formulas. By a *logic* over $\mathcal{L}$ we mean a distinguished set of formulas of $\mathcal{L}$, intuitively those regarded as valid. Such a set can be specified syntactically, by closing a collection of axioms under inference rules, or semantically, as the set of formulas satisfied by every model from a given class. When both presentations are available, soundness and completeness establish that they determine the same logic.

The basic logic underlying all formalisms considered in this thesis is classical propositional logic. Its language $\mathcal{L}_{\mathbb{P}}$ is generated from a set Atm of propositional atoms using the standard Boolean constants and connectives: verum and falsum, negation, conjunction,

disjunction, and (material) implication:

$$\mathcal{L}_{\mathbb{P}} = A ::= x \in \text{Atm} \mid \top \mid \perp \mid \neg A \mid A \wedge A \mid A \vee A \mid A \rightarrow A.$$

Semantically, formulas are evaluated with respect to Boolean valuations, which we identify with subsets $\rho \subseteq \text{Atm}$ containing exactly the atoms assigned the value true. Satisfaction of a formula $A \in \mathcal{L}_{\mathbb{P}}$ with respect to a valuation ρ is defined inductively as follows:

- $\rho \models x$ *iff* $x \in \rho$;
- $\rho \models \top$;
- $\rho \not\models \perp$;
- $\rho \models \neg A$ *iff* $\rho \not\models A$;
- $\rho \models A \wedge B$ *iff* $\rho \models A$ and $\rho \models B$;
- $\rho \models A \vee B$ *iff* $\rho \models A$ or $\rho \models B$;
- $\rho \models A \rightarrow B$ *iff* $\rho \not\models A$ or $\rho \models B$.

A formula A is *classically valid* if $\rho \models A$ for every Boolean valuation ρ . More generally, a set of formulas Γ *classically entails* A if every valuation satisfying all formulas in Γ also satisfies A . The set of all classically valid formulas constitutes classical propositional logic.

Equivalently, classical propositional logic can be characterized syntactically by a standard Hilbert calculus. For instance, taking $\neg$ and $\rightarrow$ as primitive connectives, one such calculus is given by the following axiom schemata due to Łukasiewicz [Łuk64]:

- (A1) $A \rightarrow (B \rightarrow A)$,
- (A2) $(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$,
- (A3) $(\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B)$,

together with the inference rule of *modus ponens*, which infers B from A and $A \rightarrow B$. The remaining Boolean connectives and constants can be defined from the primitive ones in the usual way. The soundness and completeness of this calculus [Łuk64] shows that the formulas obtainable from instances of these axiom schemata by finitely many applications of modus ponens are exactly the classically valid formulas (also called *classical tautologies*).

Modal logics aim to characterize a unary modality $\Box$, which can serve a variety of interpretations depending on the context (e.g., necessity, knowledge, provability). The language of modal logic is obtained by extending the language of propositional logic with the operator $\Box$:

$$\mathcal{L}_{\Box} = A ::= x \in \text{Atm} \mid \top \mid \perp \mid \neg A \mid A \wedge A \mid A \vee A \mid A \rightarrow A \mid \Box A.$$

Different modal logics capture different intended behaviours of the modality $\Box$. Among them, an important and particularly well-behaved class is that of *normal modal logics*.

The minimal normal modal logic **K** extends classical propositional logic with all instances of the axiom

$$(K) \quad \Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B),$$

and is closed under modus ponens, uniform substitution, and the *necessitation* rule, which infers $\Box A$ from A . Other normal modal logics are obtained by extending **K** with additional axiom schemata governing the behaviour of $\Box$. Axiomatisations of some normal modal logics used throughout this thesis, collected in Tab. 2.1, use the following additional modal axioms:

$$\begin{aligned} (T) \quad & \Box A \rightarrow A, \\ (4) \quad & \Box A \rightarrow \Box \Box A, \\ (5) \quad & \neg \Box A \rightarrow \Box \neg \Box A, \\ (GL) \quad & \Box(\Box A \rightarrow A) \rightarrow \Box A. \end{aligned}$$

The last axiom, known as the *Gödel–Löb axiom*, is particularly important in the study of formal provability. Under the interpretation of $\Box A$ as “ A is provable” in a sufficiently strong arithmetical theory, it reflects Löb’s theorem [Löb55]; moreover, Solovay’s arithmetical completeness theorem shows that **GL** captures exactly the modal principles of provability in Peano arithmetic [Sol76].

Normal modal logics admit a uniform semantics in terms of relational (Kripke) models [Kri66], which we recall next.

Definition 1. A relational model is a tuple $\langle W, R, V \rangle$, where W is a set of worlds, R is a binary accessibility relation on W , and $V : W \rightarrow 2^{Atm}$ is a valuation function that maps each world in W into a subset of atoms $V(w)$ satisfied in w . Satisfaction of a modal formula A in a given world $w \in W$ of a relational model $M = \langle W, R, V \rangle$ (denoted $M, w \models A$) is then defined inductively as follows:

- $M, w \models x$ if $x \in V(w)$;
- $M, w \models \top$;
- $M, w \not\models \perp$;
- $M, w \models \neg A$ if $M, w \not\models A$;
- $M, w \models A \wedge B$ if $M, w \models A$ and $M, w \models B$;
- $M, w \models A \vee B$ if $M, w \models A$ or $M, w \models B$;
- $M, w \models A \rightarrow B$ if $M, w \not\models A$ or $M, w \models B$;
- $M, w \models \Box A$ if $M, u \models A$ for all $u \in W$ s.t. $w R u$.

The truth set of a formula A w.r.t. model M is defined as $\|A\|^M = \{w : M, w \models A\}$. A is said to be valid in M if A is true in every world in M , otherwise A is said to be falsified by M .

Different normal modal logics can be defined as a set of formulas valid w.r.t. different classes of relational models. The logic of all relational models is called **K**. Tab. 2.1

Logic	Modal axioms	Class of relational models
K	(K)	all relational models
T	(K), (T)	models with reflexive R
S4	(K), (T), (4)	models with reflexive and transitive R
S5	(K), (T), (5)	models with universal R
GL	(K), (GL)	finite models with transitive and asymmetric R

Table 2.1: Axioms and frame conditions for some normal modal logics used in this thesis.

contains definitions of some other normal modal logics obtained by placing so-called *frame conditions* on the accessibility relation R . The following standard properties of the accessibility relation are used: R is *reflexive* if wRw for every $w \in W$; R is *transitive* if wRu and uRv imply wRv ; R is *asymmetric* if wRu implies that uRw does not hold; R is *universal* if wRu for all $w, u \in W$. Note that the logic **GL** is defined via *finite models*, i.e. models with finitely many worlds. All other logics in Tab. 2.1 satisfy the *finite model property* [LS77], which states that any formula falsified by a model from the corresponding class of relational models is also falsified by a finite model from the same class. So for these logics, restricting the class to the finite models does not change the resulting logic. In contrast, restricting the class of transitive and asymmetric models to finite models changes the induced logic: finite transitive asymmetric frames characterize **GL** [Seg71], whereas Löb’s axiom is not valid over arbitrary transitive asymmetric frames. So the finiteness requirement is essential to distinguish **GL**.

2.2 Sequent calculi for classical and normal modal logics

In this section, we recall the proof-theoretic framework of sequent calculi and present standard sequent calculi for classical and normal modal logics.

Sequent calculi provide a proof-theoretic framework for deriving entailment in a logic represented via *sequents*—expressions of the form $\Gamma \vdash \Delta$, where Γ and Δ are finite multisets of formulas. Under the standard interpretation, a sequent expresses that the formulas in Γ entail the disjunction of the formulas in Δ ¹. A sequent calculus is specified by a set of inference *rules*, each consisting of a (possibly empty) collection of sequents, called *premises*, and a single sequent, called the *conclusion*. Rules are typically written in the form $\frac{\mathcal{P}_1 \dots \mathcal{P}_n}{\mathcal{Q}}$, with the premises $\mathcal{P}_1, \dots, \mathcal{P}_n$ appearing above the line and the conclusion $\mathcal{Q}$ below it, where $\mathcal{P}_1, \dots, \mathcal{P}_n, \mathcal{Q}$ are meta-formulas with formula variables (denoted by capital Latin letters) in place of concrete formulas. Instantiating these formula variables with arbitrary formulas yields different *instances* of the rule. A (*closed*) *derivation* in a sequent calculus is a finite tree whose nodes are labeled with sequents, such that each node is labeled with the conclusion of some rule instance and its children are labeled with exactly all the premises. In particular, leaves correspond to instances of

¹Some applications of the sequent framework, e.g. for substructural logics, adopt different interpretations of sequents; in this thesis, we stick to the standard interpretation.

Concluding rule:

$$\frac{}{A, \Gamma \vdash A, \Delta} \text{ (id)}$$

Logical rules:

$$\begin{array}{ll} \frac{\Gamma \vdash A, \Delta}{\neg A, \Gamma \vdash \Delta} (\neg L) & \frac{A, \Gamma \vdash \Delta}{\Gamma \vdash \neg A, \Delta} (\neg R) \\ \frac{A, B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta} (\wedge L) & \frac{\Gamma \vdash A, \Delta \quad \Gamma \vdash B, \Delta}{\Gamma \vdash A \wedge B, \Delta} (\wedge R) \\ \frac{A, \Gamma \vdash \Delta \quad B, \Gamma \vdash \Delta}{A \vee B, \Gamma \vdash \Delta} (\vee L) & \frac{\Gamma \vdash A, B, \Delta}{\Gamma \vdash A \vee B, \Delta} (\vee R) \\ \frac{\Gamma \vdash A, \Delta \quad B, \Gamma \vdash \Delta}{A \rightarrow B, \Gamma \vdash \Delta} (\rightarrow L) & \frac{A, \Gamma \vdash B, \Delta}{\Gamma \vdash A \rightarrow B, \Delta} (\rightarrow R) \end{array}$$

Figure 2.1: Propositional fragment of Gentzen's sequent calculus **LK** for classical logic.

rules with zero premises (we will call such rules *concluding rules*). A sequent $\Gamma \vdash \Delta$ is called *derivable* in a calculus if there exists a closed derivation in this calculus with the root labeled with $\Gamma \vdash \Delta$.

Sequent calculi were originally introduced by Gentzen [Gen34] for (first-order) classical logic. The resulting system, known as **LK**, consists of the rules shown in Fig. 2.1: the concluding rule (id) and *logical* rules decomposing each logical connective on either side of the sequent (as well as rules for quantifiers, which we omit in our propositional setting). Derivations using only these rules are analytic, in the sense that every formula in the premises is a subformula of some formula in the conclusion (so-called *subformula property*) [Gen35]. In addition, Gentzen introduced *structural* rules (shown in Fig. 2.2), which govern the structural behavior of sequents independently of the logical connectives. The crucial structural rule (cut), which enables intermediate lemmas to be used in derivations, is not analytic however, as it allows the introduction of arbitrary formulas not present in the conclusion. Gentzen's central result (*Hauptsatz*) shows that structural rules are admissible, i.e. any sequent derivable with structural rules can also be derived without them. This allowed Gentzen to prove that **LK** is indeed sound and complete w.r.t. classical logic. Throughout the thesis, when referring to **LK**, we will mean the system consisting only of the rules in Fig. 2.1, without structural or quantifier rules.

Theorem 1 ([Gen34]). *A sequent $\Gamma \vdash \Delta$ is derivable in **LK** iff Γ classically entails $(\bigvee_{A \in \Delta} A)$.*

Later, the sequent calculus framework was successfully extended to normal modal logics. Sequent calculi for such logics are typically obtained by extending **LK** with rules for eliminating the $\Box$ modality. Examples of such rules for some of the normal modal logics

$$\frac{\Gamma \vdash \Delta}{A, \Gamma \vdash \Delta} \text{ (wk)} \quad \frac{A, A, \Gamma \vdash \Delta}{A, \Gamma \vdash \Delta} \text{ (ctr)} \quad \frac{\Gamma \vdash A, \Delta \quad A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta} \text{ (cut)}$$

Figure 2.2: Structural sequent rules

$$\frac{\Gamma^{\Box\downarrow} \vdash A}{\Gamma \vdash \Box A, \Delta} (\Box_K R) \quad \frac{A, \Gamma \vdash \Delta}{\Box A, \Gamma \vdash \Delta} (\Box_T L) \quad \frac{\Gamma^{\Box}, \Gamma^{\Box\downarrow} \vdash A}{\Gamma \vdash \Box A, \Delta} (\Box_4 R) \quad \frac{\Gamma^{\Box}, \Gamma^{\Box\downarrow}, \Box A \vdash A}{\Gamma \vdash \Box A, \Delta} (GL)$$

$$\begin{aligned} SC_K &= \mathbf{LK} \cup \{(\Box_K R)\} & SC_T &= \mathbf{LK} \cup \{(\Box_K R), (\Box_T L)\} \\ SC_{GL} &= \mathbf{LK} \cup \{(GL)\} & SC_{S4} &= \mathbf{LK} \cup \{(\Box_4 R), (\Box_T L)\} \end{aligned}$$

Figure 2.3: Sequent calculi for some normal modal logics.

introduced in the previous section are shown in Fig. 2.3. They use the auxiliary notation $\Gamma^{\Box} = \{\Box A \mid \Box A \in \Gamma\}$ and $\Gamma^{\Box\downarrow} = \{A \mid \Box A \in \Gamma\}$, denoting, respectively, the submultiset of boxed formulas occurring in Γ and the multiset obtained by removing the outermost $\Box$ from each of them. In particular, Valentini’s calculus for **GL** [Val83] extends **LK** with a single “diagonal” rule (GL), which removes $\Box$ from a formula on the RHS of a sequent while simultaneously moving this boxed formula to the LHS. Although the premise of this rule contains more formulas than the conclusion, Valentini showed that it is “simpler” with respect to some natural measure of complexity on sequents, resulting in an elegant analytic calculus with terminating proof search for this non-trivial logic.

Theorem 2 ([Pog11, SV82]). *For every logic $\mathcal{L} \in \{\mathbf{K}, \mathbf{T}, \mathbf{S4}, \mathbf{GL}\}$, a sequent $\Gamma \vdash \Delta$ is derivable in $SC_{\mathcal{L}}$ (as defined in Fig. 2.3) iff $(\bigwedge_{A \in \Gamma} A) \rightarrow (\bigvee_{B \in \Delta} B)$ is valid in $\mathcal{L}$.*

In contrast, the logic **S5** resists such a direct treatment within the sequent framework: no standard cut-free sequent calculus is known for it, and analytic calculi for **S5** require more expressive structures in derivations beyond standard sequents.

2.3 Hypersequent calculi

While the standard sequent framework works well for many normal modal logics such as **K**, **T**, and **S4**, it cannot capture all modal systems without using the cut rule. The main limitation is a structural mismatch between standard sequent derivations and certain frame properties of relational models. Sequent rules operate locally, propagating information forward along the branches of a proof tree. This behaviour mirrors reflexive and transitive accessibility relations, but when symmetry is added to the semantics—as in the case of **S5**—this local, forward-moving style of reasoning is no longer sufficient. As a result, no cut-free sequent calculus for **S5** is known.

A standard way to overcome this limitation is to generalize sequents to *hypersequents*, finite collections of ordinary sequents interpreted disjunctively. By allowing several sequents to be manipulated simultaneously, hypersequents provide enough structure

Concluding rule:

$$\frac{}{A, \Gamma \vdash A, \Delta \mid \mathcal{H}} (id^{\mathcal{H}})$$

Logical rules:

$$\begin{array}{c} \frac{\Gamma \vdash A, \Delta \mid \mathcal{H}}{\neg A, \Gamma \vdash \Delta \mid \mathcal{H}} (\neg L^{\mathcal{H}}) \qquad \frac{A, \Gamma \vdash \Delta \mid \mathcal{H}}{\Gamma \vdash \neg A, \Delta \mid \mathcal{H}} (\neg R^{\mathcal{H}}) \\[10pt] \frac{A, B, \Gamma \vdash \Delta \mid \mathcal{H}}{A \wedge B, \Gamma \vdash \Delta \mid \mathcal{H}} (\wedge L^{\mathcal{H}}) \qquad \frac{\Gamma \vdash A, \Delta \mid \mathcal{H} \quad \Gamma \vdash B, \Delta \mid \mathcal{H}}{\Gamma \vdash A \wedge B, \Delta \mid \mathcal{H}} (\wedge R^{\mathcal{H}}) \\[10pt] \frac{A, \Gamma \vdash \Delta \mid \mathcal{H} \quad B, \Gamma \vdash \Delta \mid \mathcal{H}}{A \vee B, \Gamma \vdash \Delta \mid \mathcal{H}} (\vee L^{\mathcal{H}}) \qquad \frac{\Gamma \vdash A, B, \Delta \mid \mathcal{H}}{\Gamma \vdash A \vee B, \Delta \mid \mathcal{H}} (\vee R^{\mathcal{H}}) \\[10pt] \frac{\Gamma \vdash A, \Delta \mid \mathcal{H} \quad B, \Gamma \vdash \Delta \mid \mathcal{H}}{A \rightarrow B, \Gamma \vdash \Delta \mid \mathcal{H}} (\rightarrow L^{\mathcal{H}}) \qquad \frac{A, \Gamma \vdash B, \Delta \mid \mathcal{H}}{\Gamma \vdash A \rightarrow B, \Delta \mid \mathcal{H}} (\rightarrow R^{\mathcal{H}}) \end{array}$$

Modal rules:

$$\frac{A, \Gamma \vdash \Delta \mid \mathcal{H}}{\Box A, \Gamma \vdash \Delta \mid \mathcal{H}} (\Box_{\mathbf{S5}} L^{\mathcal{H}}) \qquad \frac{\Gamma^{\Box}, \Gamma^{\Box\downarrow} \vdash A \mid \mathcal{H}}{\Gamma \vdash \Box A, \Delta \mid \mathcal{H}} (\Box_{\mathbf{S5}} R^{\mathcal{H}}) \qquad \frac{\Gamma^{\Box}, \Sigma \vdash \Pi \mid \mathcal{H}}{\Gamma \vdash \Delta \mid \Sigma \vdash \Pi \mid \mathcal{H}} (com_{\mathbf{S5}}^{\mathcal{H}})$$

Structural rules:

$$\begin{array}{c} \frac{\Gamma \vdash \Delta \mid \mathcal{H}}{A, \Gamma \vdash \Delta \mid \mathcal{H}} (iwk^{\mathcal{H}}) \qquad \frac{A, A, \Gamma \vdash \Delta \mid \mathcal{H}}{A, \Gamma \vdash \Delta \mid \mathcal{H}} (ict^{\mathcal{H}}) \\[10pt] \frac{\mathcal{H}}{\Gamma \vdash \Delta \mid \mathcal{H}} (ewk^{\mathcal{H}}) \qquad \frac{\Gamma \vdash \Delta \mid \Gamma \vdash \Delta \mid \mathcal{H}}{\Gamma \vdash \Delta \mid \mathcal{H}} (ect^{\mathcal{H}}) \end{array}$$

Figure 2.4: Hypersequent calculus $\mathcal{H}_{\mathbf{S5}}$ for $\mathbf{S5}$.

to capture interactions between multiple worlds that cannot be represented within a single sequent. Originally introduced precisely to obtain analytic proof systems for **S5** [Pot83], the hypersequent framework has since proved useful for other systems resisting standard sequent treatment as well, including modal, intermediate, substructural and fuzzy logics [Avr96b, CGT08, MOG08, Lah13]. In this section, we introduce the hypersequent formalism and the hypersequent calculus for **S5**, which will later serve as a building block for our proof-theoretic developments for conditional logics.

A hypersequent has the form

$$\Gamma_1 \vdash \Delta_1 \mid \cdots \mid \Gamma_n \vdash \Delta_n,$$

where Γ_i and Δ_i are multisets of modal formulas. Each $\Gamma_i \vdash \Delta_i$ is called a *component* of such hypersequent. In the **S5** setting considered here, the hypersequent can be interpreted as the formula

$$\bigvee_{1 \leq i \leq n} \Box((\bigwedge_{A \in \Gamma_i} A) \rightarrow (\bigvee_{B \in \Delta_i} B)).$$

Inference rules have hypersequents both as premises and conclusions.

A hypersequent calculus $\mathcal{H}_{\mathbf{S5}}$ for **S5** (adapted from [Kur13]) is presented in Fig. 2.4 (see [Avr96a] for a different calculus). It consists of the hypersequent versions of the rules of the sequent calculus $SC_{\mathbf{S4}}$ for **S4** (Fig. 2.3) together with structural rules operating on whole components, and an additional *communication rule* ($com_{\mathbf{S5}}^{\mathcal{H}}$). The hypersequent versions of the rules of $SC_{\mathbf{S4}}$ is simply obtained by adding a hypersequent context $\mathcal{H}$ (representing a possibly empty hypersequent) to each sequent in the calculus. Structural rules in hypersequent calculi are divided into two classes: *internal* rules, which operate within individual components, and *external* rules. Examples of the latter are the rule ($ectr^{\mathcal{H}}$), which duplicates components, and the rule ($ewk^{\mathcal{H}}$) which weakens a whole component.

Recall that a rule is *admissible* in a calculus if, whenever all premises of an instance of the rule are derivable, its conclusion is also derivable without using that rule; equivalently, adding the rule does not change the set of derivable hypersequents. In $\mathcal{H}_{\mathbf{S5}}$, the internal structural rules are admissible, just as in the sequent calculi considered previously. By contrast, external contraction is not admissible in this presentation and must therefore be included explicitly in order to obtain a complete calculus (see Ex. 1 below). The presence of multiple sequents in hypersequents allows one to add new external structural rules which increase the expressive power of the hypersequent framework w.r.t. the sequent framework. The peculiar rule for the logic **S5** is the communication rule ($com_{\mathbf{S5}}^{\mathcal{H}}$) which acts on two components simultaneously and allows boxed formulas occurring on the left of one component to be transferred to the other. This reflects the universality of the accessibility relation in **S5**: since every world can access every other world, boxed formulas that are true in some world are also true in all the others. Together, these rules constitute a sound and complete proof system for **S5**.

Theorem 3 ([Kur13]). *A hypersequent $\Gamma_1 \vdash \Delta_1 \mid \cdots \mid \Gamma_n \vdash \Delta_n$ is derivable in $\mathcal{H}_{\mathbf{S5}}$ iff the formula $\bigvee_{1 \leq i \leq n} \Box((\bigwedge_{A \in \Gamma_i} A) \rightarrow (\bigwedge_{B \in \Delta_i} B))$ is valid in **S5**.*

Concluding rule:

$$\frac{}{A, \Gamma \vdash A, \Delta \mid \mathcal{H}} (id^K)$$

Logical rules:

$$\begin{array}{c} \frac{\neg A, \Gamma \vdash \boxed{A}, \Delta \mid \mathcal{H}}{\neg A, \Gamma \vdash \Delta \mid \mathcal{H}} (\neg L^K) \quad \frac{\boxed{A}, \Gamma \vdash \neg A, \Delta \mid \mathcal{H}}{\Gamma \vdash \neg A, \Delta \mid \mathcal{H}} (\neg R^K) \\[10pt] \frac{\boxed{A, B}, A \wedge B, \Gamma \vdash \Delta \mid \mathcal{H}}{A \wedge B, \Gamma \vdash \Delta \mid \mathcal{H}} (\wedge L^K) \quad \frac{\Gamma \vdash \boxed{A}, A \wedge B, \Delta \mid \mathcal{H} \quad \Gamma \vdash \boxed{B}, A \wedge B, \Delta \mid \mathcal{H}}{\Gamma \vdash A \wedge B, \Delta \mid \mathcal{H}} (\wedge R^K) \\[10pt] \frac{\boxed{A}, A \vee B, \Gamma \vdash \Delta \mid \mathcal{H} \quad \boxed{B}, A \vee B, \Gamma \vdash \Delta \mid \mathcal{H}}{A \vee B, \Gamma \vdash \Delta \mid \mathcal{H}} (\vee L^K) \quad \frac{\Gamma \vdash \boxed{A, B}, A \vee B, \Delta \mid \mathcal{H}}{\Gamma \vdash A \vee B, \Delta \mid \mathcal{H}} (\vee R^K) \\[10pt] \frac{A \rightarrow B, \Gamma \vdash \boxed{A}, \Delta \mid \mathcal{H} \quad \boxed{B}, A \rightarrow B, \Gamma \vdash \Delta \mid \mathcal{H}}{A \rightarrow B, \Gamma \vdash \Delta \mid \mathcal{H}} (\rightarrow L^K) \quad \frac{\boxed{A}, \Gamma \vdash \boxed{B}, A \rightarrow B, \Delta \mid \mathcal{H}}{\Gamma \vdash A \rightarrow B, \Delta \mid \mathcal{H}} (\rightarrow R^K) \end{array}$$

Modal rules:

$$\begin{array}{c} \frac{\boxed{A}, \Box A, \Gamma \vdash \Delta \mid \mathcal{H}}{\Box A, \Gamma \vdash \Delta \mid \mathcal{H}} (\Box_{S5} L^K) \quad \frac{\boxed{\Gamma^\Box, \Gamma^{\Box\downarrow} \vdash A \mid \Gamma \vdash \Box A, \Delta \mid \mathcal{H}}}{\Gamma \vdash \Box A, \Delta \mid \mathcal{H}} (\Box_{S5} R^K) \\[10pt] \frac{\Gamma \vdash \Delta \mid \boxed{\Gamma^\Box}, \Sigma \vdash \Pi \mid \mathcal{H}}{\Gamma \vdash \Delta \mid \Sigma \vdash \Pi \mid \mathcal{H}} (com_{S5}^K) \end{array}$$

Figure 2.5: Kleene-style hypersequent calculus $\mathcal{HK}_{S5}$ for **S5**. The addition in each premise w.r.t. the conclusion is highlighted.

Notice that this generalized hypersequent formulation of soundness and completeness subsumes the standard sequent formulation as a special case.

Corollary 1. *A single-component hypersequent $\Gamma \vdash \Delta$ is derivable in $\mathcal{H}_{S5}$ iff Γ entails $(\bigvee_{A \in \Delta} A)$ in **S5**.*

Example 1. *Axiom (5) of **S5** can be derived in $\mathcal{H}_{S5}$ as follows:*

$$\begin{array}{c} \frac{}{\Box A \vdash \Box A, \Box \neg \Box A} (id^H) \\ \frac{}{\Box A \vdash \vdash \Box A, \Box \neg \Box A} (com_{S5}^H) \\ \frac{}{\vdash \neg \Box A \mid \vdash \Box A, \Box \neg \Box A} (\neg R^H) \\ \frac{}{\vdash \Box A, \Box \neg \Box A \mid \vdash \Box A, \Box \neg \Box A} (\Box_{S5} R^H) \\ \frac{}{\vdash \Box A, \Box \neg \Box A} (ectr^H) \\ \frac{}{\neg \Box A \vdash \Box \neg \Box A} (\neg L^H) \end{array}$$

Notice that external contraction is essential here: without applying $(ctr^{\mathcal{H}})$, the endsequent is not derivable, since after applying $(\neg L^{\mathcal{H}})$ the only productive rule applications of $(\Box_{\mathbf{S5}} R^{\mathcal{H}})$ eliminate one of the boxed formulas on the right and lead to an underivable premise.

In this thesis, we will employ a particular style of hypersequent calculi known as *Kleene-style calculi* [KL18]. In rules in such calculi, each premise repeats the entire conclusion and extends it by adding new formulas or new components. This makes all rules *invertible*, meaning that premises of any rule application remain derivable if the conclusion was. Consequently, when derivations are read bottom-up, information accumulates monotonically. In proof search, we consider only *productive* rule applications, i.e. applications that add a formula or component not already present in the hypersequent. Since only subformulas of the end-hypersequent can be introduced, there are only finitely many distinct formulas and components that can occur. Hence, after finitely many steps, every branch either can be closed or terminates in a finite hypersequent to which no productive rule is applicable; we call such a hypersequent *saturated*.

Calculus $\mathcal{HK}_{\mathbf{S5}}$, the Kleene-style version of $\mathcal{H}_{\mathbf{S5}}$, is presented in Fig. 2.5. The rules are obtained by retaining the principal formula in the premises. In all rules except $(\Box_{\mathbf{S5}} R^{\mathcal{K}})$, the additional formulas introduced by the rule are added directly to the principal component. The rule $(\Box_{\mathbf{S5}} R^{\mathcal{K}})$ is different: semantically, the modified component corresponds to a potentially different world of the model, therefore it is added to hypersequent as a separate component. Unlike $\mathcal{H}_{\mathbf{S5}}$, Kleene-style calculus $\mathcal{HK}_{\mathbf{S5}}$ contains no explicit structural rules, since these become admissible in Kleene form.

Lemma 1. *The structural rules from Fig. 2.4 are admissible in $\mathcal{HK}_{\mathbf{S5}}$.*

Proof. For weakening, this follows directly from the form of the rules: additional formulas within components and additional components can be retained unchanged in both conclusions and premises throughout a derivation. For contraction, consider derivations bottom-up. Since every Kleene-style rule preserves the entire conclusion in its premises and only adds formulas or components, any rule application performed using one of two identical occurrences of a formula, or one of two identical components, can instead be performed using the other copy. Thus, after identifying the two copies, all subsequent rule applications can be performed on the single remaining occurrence or component (and an application of $(com_{\mathbf{S5}}^{\mathcal{K}})$ between the two identified copies then becomes redundant), without preventing any branch from closing. Hence applications of both internal and external contraction can be eliminated. $\square$

It is straightforward to translate derivations between $\mathcal{H}_{\mathbf{S5}}$ and $\mathcal{HK}_{\mathbf{S5}}$. Thus, the calculi $\mathcal{H}_{\mathbf{S5}}$ and $\mathcal{HK}_{\mathbf{S5}}$ are equivalent with respect to derivability, and both capture **S5**.

Theorem 4. *A hypersequent $\Gamma_1 \vdash \Delta_1 \mid \dots \mid \Gamma_n \vdash \Delta_n$ is derivable in $\mathcal{HK}_{\mathbf{S5}}$ iff the formula $\bigvee_{1 \leq i \leq n} \Box((\bigwedge_{A \in \Gamma_i} A) \rightarrow (\bigwedge_{B \in \Delta_i} B))$ is valid in **S5**.*

Proof. By Th. 3, it suffices to show that $\mathcal{HK}_{\mathbf{S5}}$ and $\mathcal{H}_{\mathbf{S5}}$ derive the same hypersequents.

First, every $\mathcal{HK}_{\mathbf{S5}}$ -derivation can be translated into an $\mathcal{H}_{\mathbf{S5}}$ -derivation. Each application of a Kleene-style rule in $\mathcal{HK}_{\mathbf{S5}}$ can be simulated in $\mathcal{H}_{\mathbf{S5}}$ by first applying, bottom-up, suitable instances of internal or external contraction, followed by the corresponding rule of $\mathcal{H}_{\mathbf{S5}}$.

Conversely, every $\mathcal{H}_{\mathbf{S5}}$ -derivation can be translated into an $\mathcal{HK}_{\mathbf{S5}}$ -derivation by induction on its height. Applications of structural rules are eliminated using their admissibility in $\mathcal{H}_{\mathbf{S5}}$. Every remaining rule is simulated by its Kleene-style counterpart, using admissible weakening to add the principal formulas or components retained in its premises.

Thus, a hypersequent is derivable in $\mathcal{HK}_{\mathbf{S5}}$ iff it is derivable in $\mathcal{H}_{\mathbf{S5}}$, and the claim follows from Th. 3. $\square$

Example 2. *Axiom (5) of $\mathbf{S5}$ can be derived in $\mathcal{HK}_{\mathbf{S5}}$ as follows (cf. derivation in $\mathcal{H}_{\mathbf{S5}}$ in Ex. 1):*

$$\begin{array}{c}
 \frac{}{\Box A \vdash \neg \Box A \mid \Box A, \neg \Box A, \vdash \Box A, \Box \neg \Box A} (id^{\mathcal{H}}) \\
 \frac{}{\Box A \vdash \neg \Box A \mid \neg \Box A, \vdash \Box A, \Box \neg \Box A} (com_{\mathbf{S5}}^{\mathcal{H}}) \\
 \frac{}{\vdash \neg \Box A \mid \neg \Box A \vdash \Box A, \Box \neg \Box A} (\neg R^{\mathcal{H}}) \\
 \frac{}{\neg \Box A \vdash \Box A, \Box \neg \Box A} (\Box_{\mathbf{S5}} R^{\mathcal{K}}) \\
 \frac{}{\neg \Box A \vdash \Box \neg \Box A} (\neg L^{\mathcal{K}})
 \end{array}$$

The Kleene form of the rules is particularly well suited for semantic analysis of the calculus, since saturated hypersequents often contain sufficient information to reconstruct countermodels directly, providing a neat semantical way to establish completeness of the calculus. For this reason, when defining hypersequent calculi for conditional logics (in Chapters 4 and 7) we will use the Kleene form of the rules.

2.4 Antisequent calculus

Traditionally, proof theory is concerned with formal systems for deriving valid formulas and valid entailments. However, there is also a complementary line of research devoted to deriving *non-validity*. This task is particularly important in formalisms where validity and non-validity are intrinsically intertwined and defined in terms of each other.

A simple proof-theoretic formalism of this kind for classical logic was introduced in [Bon93, BO02] under the name of *antisequent calculus*. It operates with *antisequents* of the form $\Gamma \not\vdash \Delta$, where Γ and Δ are multisets of propositional formulas. The intended interpretation is dual to that of ordinary sequents: an antisequent $\Gamma \not\vdash \Delta$ is valid when $\bigwedge_{A \in \Gamma} A$ does *not* classically imply $\bigvee_{B \in \Delta} B$. The rules of the calculus *antiLK* (Fig. 2.6) consist of elimination rules for propositional connectives, dual to the corresponding decomposition rules of **LK**, together with a concluding rule requiring that Γ and Δ are disjoint. This mirrors the identity axiom of **LK**, where the presence of a common formula on both sides

Concluding rule:

$$\frac{}{\Gamma \not\vdash \Delta} (\not\vdash id) \quad \text{when } \Gamma \cup \Delta \subseteq \text{Atm} \text{ and } \Gamma \cap \Delta = \emptyset$$

Logical rules:

$$\begin{array}{c} \frac{\Gamma \not\vdash A, \Delta}{\neg A, \Gamma \not\vdash \Delta} (\not\vdash \neg L) \quad \frac{A, \Gamma \not\vdash \Delta}{\Gamma \not\vdash \neg A, \Delta} (\not\vdash \neg R) \\[10pt] \frac{A, B, \Gamma \not\vdash \Delta}{A \wedge B, \Gamma \not\vdash \Delta} (\not\vdash \wedge L) \quad \frac{\Gamma \not\vdash A, \Delta}{\Gamma \not\vdash A \wedge B, \Delta} (\not\vdash \wedge R_1) \quad \frac{\Gamma \not\vdash B, \Delta}{\Gamma \not\vdash A \wedge B, \Delta} (\not\vdash \wedge R_2) \\[10pt] \frac{A, \Gamma \not\vdash \Delta}{A \vee B, \Gamma \not\vdash \Delta} (\not\vdash \vee L_1) \quad \frac{B, \Gamma \not\vdash \Delta}{A \vee B, \Gamma \not\vdash \Delta} (\not\vdash \vee L_2) \quad \frac{\Gamma \not\vdash A, B, \Delta}{\Gamma \not\vdash A \vee B, \Delta} (\not\vdash \vee R) \\[10pt] \frac{\Gamma \not\vdash A, \Delta}{A \rightarrow B, \Gamma \not\vdash \Delta} (\not\vdash \rightarrow L_1) \quad \frac{B, \Gamma \not\vdash \Delta}{A \rightarrow B, \Gamma \not\vdash \Delta} (\not\vdash \rightarrow L_2) \quad \frac{A, \Gamma \not\vdash B, \Delta}{\Gamma \not\vdash A \rightarrow B, \Delta} (\not\vdash \rightarrow R) \end{array}$$

Figure 2.6: Antisequent calculus *antiLK* for the classical non-validity

closes a branch. Unlike in **LK**, however, the concluding rule can only be applied once all logical connectives have been eliminated and the antisequent contains only propositional atoms, otherwise same formula can still appear on both sides as a result of further eliminations.

Theorem 5 ([Bon93]). *The antisequent $\Gamma \not\vdash \Delta$ is derivable in *antiLK* iff Γ does not classically entail $(\bigvee_{A \in \Delta} A)$.*

Antisequent calculus has subsequently been employed as a building block in the proof-theoretic study of several globally non-monotonic logical formalisms, including circumscription, default and autoepistemic logics [Bon96, BO97, BO98, BO02]. In Chapter 7, we will use it in a similar way as one of the ingredients of our analytic calculus for rational closure.

Part I

Counterfactuals

Preferential Conditional Logic

This chapter studies the counterfactual interpretation of conditionals, where a formula $A \rightsquigarrow B$ is read as “if A were the case, then B would be the case,” capturing reasoning about hypothetical situations, often contrary to the actual facts. As the formal framework for such reasoning, we consider Preferential Conditional Logic (**PCL**) of Burgess [Bur81], which is based on semantics of *preferential models* equipped with similarity orderings between possible worlds. **PCL** occupies an important position both within philosophical logic, where it provides an abstract generalization of Lewis-style similarity semantics [Lew73], and it is also closely connected to preferential approach for non-monotonic reasoning in symbolic AI [Sho87], making it an important meeting point between philosophical and computational approaches to conditional reasoning.

We introduce a simple, natural and complexity-optimal analytic sequent calculus for **PCL**. The calculus is obtained by extending the approach proposed in [GGOP09] for the KLM logic **P**, corresponding to the flat fragment of **PCL**. That approach exploits a connection between **P** and Gödel–Löb provability logic **GL**. To extend this connection to the nested setting, we rely on the analytic semantic characterization of **PCL** provided by the small-model constructions of Friedman and Halpern [FH94]. We begin by surveying the historical background (Section 3.1) and the state of the art regarding analytic methods (Section 3.2). We then formally introduce preferential semantics and **PCL** (Section 3.3). In Section 3.4, we present a new sequent calculus for full **PCL**. The calculus combines complexity-optimal proof search with reconstruction of countermodels from failed derivations, thereby providing a constructive way to obtain the Friedman–Halpern model constructions. In Sec. 3.5, we formalize the connection between **PCL** and provability logics in the form of an embedding of **PCL** into the bimodal logic **GL+K** and use our calculus to establish its adequacy.

3.1 Background

Classical material implication is not suitable for reasoning about counterfactual conditionals: it renders a conditional automatically true whenever the antecedent is false, failing to distinguish between meaningful counterfactuals (e.g., “If it had rained earlier today, the grass would be wet.”) and nonsensical ones (e.g., “If it had rained earlier today, all cats would be speaking Latin.”). Early attempts to refine this, such as C.I. Lewis’s notion of strict implication [LL32], were also too rigid: they required a necessary logical connection between A and B that is rarely present in empirical or causal reasoning. The technical challenge was clarified in the seminal paper by Goodman [Goo47], who pointed out that the truth of a counterfactual depends on which background facts should remain constant and which should change in a counterfactual scenario, a distinction that seems impossible to make via linguistic rules alone without falling into a vicious circle. The problem was eventually solved via the application of a formal possible-world semantics. It was first formalized by Stalnaker [Sta68] by introducing a notion of “selection function”, picking the possible world satisfying the antecedent that is “closest” to the current world. This approach, however, relied on the controversial assumption that a unique closest world always exists. Lewis refined this framework in his landmark book *Counterfactuals* [Lew73], replacing the selection function with a system of nested ‘spheres’ around each world, encoding a graded measure of similarity between worlds. Lewis’s similarity-based framework had a profound impact on analytic philosophy, establishing counterfactuals as the primary tool for the contemporary analysis of causation [Lew86], scientific explanation [Woo03], and the laws of nature [Lew79]. Lewis’s possible-world approach also found early applications in Artificial Intelligence: Ginsberg [Gin85] adapted it to formalize counterfactual reasoning for AI, discussing applications to planning, design, and automated diagnosis.

Lewis also demonstrated that his sphere systems could be equivalently expressed through a comparative similarity relation between worlds. Exploring this idea further, Burgess shifted the perspective toward studying such possible-world models with an ordering relation as independent mathematical structures, introducing the framework [Bur81] now known as **PCL** (for Preferential Conditional Logic, sometimes also referred to as system $\mathcal{S}$ or system **B**). In this generalized framework, he replaced Lewis’s total similarity orderings with abstract partial orders, accommodating incomparable worlds, and obtaining an elegant formalism directly connected to relational models in modal logic. Preferential models since became a unifying tool in the study of conditionals, extending their influence well beyond the counterfactual tradition, e.g. to normative reasoning (see Chapter 4) and defeasible reasoning (see Chapters 6 and 7).

3.2 State-of-the-art

The formalization of **PCL** as a mathematical framework naturally led to a search for effective decision procedures and analytic characterizations. On the semantic side, a seminal paper by Friedman and Halpern [FH94] introduced small-model constructions

for **PCL** and its extensions as a tool for analyzing computational complexity. They showed that any preferential model satisfying a formula can be transformed into another model of bounded size satisfying the same formula, and used this result to establish PSPACE-completeness of **PCL**. Their construction is analytic in the sense that it proceeds inductively along the structure of the formula; however, it does not directly yield an explicit method for building satisfying models from satisfiable formulas, as their decision procedure relies on guessing suitable assignments of propositional variables.

Analytic proof-theoretic characterizations appeared later, and while several calculi have been proposed, they typically have rather complex structure and face a trade-off between computational optimality and a transparent connection to the semantics. To our knowledge, the only complexity-optimal proof system is the tableau calculus in [SPH10], which employs a complicated combinatorial rule enumerating possible orderings of conditionals, leading to exponential branching in derivations, and relies on sophisticated syntactic arguments for completeness, thus lacking a direct semantic interpretation. Other proof-theoretic approaches maintain a closer link to semantics at the cost of suboptimal computational behavior and some additional technical overhead: labeled tableaux and sequent calculi [GGOS09, GNO21, NO15] rely on richer syntactic machinery of labels, the resolution-based approach [NP18] depends on non-trivial pre-processing of formulas into normal form.

At the same time, a more lightweight proof-theoretic treatment has been obtained for a logic **P** used in reasoning about typicality (discussed in detail in Chapter 6), which coincides with the flat fragment of **PCL** (i.e. a fragment that prohibits nesting of conditionals) used in reasoning about typicality. Giordano et al. [GGOP09] introduced a tableau calculus for **P**, based on a connection with Gödel–Löb provability logic **GL**. **GL**, a normal modal logic capturing formal provability in arithmetic, admits a well-known simple and elegant proof calculus due to Valentini [SV82] (see Sec. 2.2 for details). Exploiting the correspondence between preference relations in preferential models and accessibility relations in **GL**, Giordano et al. constructed an embedding of **P** into **GL** and develop a tableau calculus by enriching the language with a modality for preference relation and employing Valentini’s rule for it. Extending this approach to full **PCL**, however, proves challenging: subsequent work [GGOS09] adapting it to **PCL** employed additional machinery of labels and intricate blocking conditions to ensure termination, while also losing a direct formal connection with **GL**.

3.3 Preliminaries: Preferential models and **PCL**

Conditional logics deal with conditionals represented by a binary modality (in the counterfactual setting, typically denoted $>$). The language of conditional logic **PCL** is therefore given as follows:

$$\mathcal{L}_{\mathbf{PCL}} = A ::= x \in \text{Atm} \mid \neg A \mid A \wedge A \mid A \vee A \mid A \rightarrow A \mid A > A.$$

PCL is defined in terms of generalized preferential models of Burgess [Bur81], which can be seen as a variation of relational models where, for each world w , the set of worlds accessible from w is equipped with a partial order. In the counterfactual setting, this partial order is understood as comparing worlds by their “closeness” to w .

Definition 2. A (counterfactual) preferential model is a tuple $\langle W, R, \{\succeq_w\}_{w \in W}, V \rangle$, where W is a set of worlds, R is a binary accessibility relation on W , $\succeq_w$ is a partial order on $R(w) = \{u \in W : w R u\}$, and $V : W \rightarrow 2^{Atm}$ is a valuation function. The strict version of $\succeq_w$, is defined as follows: $v \succ_w u$ when $v \succeq_w u$ and $u \not\succeq_w v$. Satisfaction of a modal formula A in a given world $w \in W$ of a preferential model $M = \langle W, R, \{\succeq_w\}_{w \in W}, V \rangle$ is defined inductively, for atoms and Boolean connectives in the same way as for relational models, and for conditionals as follows:

$$M, w \models A > B \text{ if } \begin{array}{l} \text{for every } u \in R(w) \text{ if } M, u \models A \text{ then there exists } v \in R(w) \\ \text{s.t. } v \succeq_w u \text{ and } M, v \models A \text{ and for every } v' \in R(w), \\ \text{if } v' \succeq_w v \text{ then } M, v' \models A \rightarrow B. \end{array}$$

This complicated definition of the satisfaction of a conditional is necessary to deal with the general case, where there may be no “closest” worlds w.r.t. $\succeq_w$. In the restricted case where $\succeq_w$ for every $w \in W$ is well-founded (i.e. there is no infinite ascending chain $u_0 \prec_w u_1 \prec_w \dots$ in M), this definition is equivalent to the fact that B is satisfied in all closest worlds, among those that satisfy A , as demonstrated in the following folklore result.

Lemma 2. For a preferential model $M = \langle W, R, \{\succeq_w\}_{w \in W}, V \rangle$, if $\succeq_w$ is well-founded for every $w \in W$ then $M, w \models A > B$ iff $M, u \models B$ for every $u \in \max_{\succ_w}(\|A\|^M \cap R(w))$, where $\max_{\succ_w}(U) = \{u \in U : \nexists v \in U \text{ s.t. } v \succ_w u\}$.

Proof. Left-to-right: If $u \in \max_w(\|A\|^M \cap R(w))$, we can apply truth condition from Def. 2 to u and get $v \in R(w)$ s.t. $M, v \models A$ and $v \succeq_w u$ (and, therefore, $u \succeq_w v$ due to maximality of u) and for every $v' \in R(w)$, $v' \succeq_w v$ implies $M, v' \models A \rightarrow B$. The latter also applies to u , so $M, u \models B$.

Right-to-left: For every $u \in \|A\|^M \cap R(w)$ there exists $v \in \max_w(\|A\|^M \cap R(w))$ such that $v \succeq_w u$ (any $\succ_w$ -ascending chain in $\|A\|^M \cap R(w)$ starting from u ends with such a world due to well-foundedness of $\succeq_w$), and for every $v' \in R(w)$ if $v' \succeq_w v$ and $M, v' \models A$ then $v' \in \max_w(\|A\|^M \cap R(w))$ due to maximality of v and transitivity of $\succeq_w$, therefore $M, v' \models B$. $\square$

While the well-foundedness assumption (also known as “limit assumption”) is not desirable for the general class of preferential models for conceptual reasons, since our analytic account will focus on finite (and therefore trivially well-founded) models, we can rely on this simpler characterization of the satisfaction of a conditional.

3.4 GL-based sequent calculus for PCL

The starting point for our calculus for **PCL** is the translation into **GL** and the corresponding analytic (tableaux) calculus developed by Giordano et al. [GGOP09] for the logic **P**, which coincides with the flat fragment of **PCL**. In this approach, an additional modality $\Box$ is introduced, where $\Box A$ is true at a world whenever A holds in all worlds strictly preferred to it with respect to the underlying preference relation. In the case of **P**, all conditionals are evaluated at the same world, and hence only a single preference relation is relevant for the evaluation (in fact, preferential models for **P** are standardly defined with a single global preference relation). Giordano et al. show that a conditional $A > B$ can be equivalently expressed using the $\Box$ modality as

$$((A \wedge \Box \neg A) \rightarrow B) \wedge \Box((A \wedge \Box \neg A) \rightarrow B).$$

Moreover, they observe that the $\Box$ modality can be treated analogously to the $\Box$ modality in logic **GL**, since **P** (and more generally **PCL**) enjoys the finite model property and the strict preference relation satisfies the same frame conditions as the accessibility relation in **GL**. This correspondence is then used both to propose an embedding of **P** into **GL** and to develop a proof calculus for **P** by employing the elimination rule (*GL*) of Valentini to the $\Box$ modality.

However, this approach does not extend directly to **PCL**. Due to the presence of nested conditionals, different conditionals may be evaluated at different worlds, and thus with respect to different preference relations; using a single $\Box$ -modality to represent all of them leads to semantically unsound reasoning. To address this issue, we develop an alternative translation of **PCL** into the bimodal logic **GL+K**.

The intuition behind our translation is inspired by the small-model constructions for **PCL** due to Friedman and Halpern [FH94]. In their approach, a model satisfying a given formula is built as a tree-like structure of *clusters* of worlds equipped with an internal ordering, where each child-cluster ensures the correct evaluation of certain conditionals at a world in its parent-cluster. Since each cluster is a finite submodel with a transitive ordering, it can be viewed as a separate **GL**-submodel, while the tree structure connecting clusters can be captured by an additional accessibility relation without frame conditions, corresponding to logic **K**. This suggests extending the translation of a conditional in **P** into **GL** by adding an independent **K**-modality $\Box_K$ on top of it. Furthermore, we observe that the translation for **P** above can be simplified by omitting the first conjunct. This simplification relies on a countermodel transformation that introduces a “dummy” world strictly less preferred than the falsifying world in the original countermodel (formal details are provided in the proof of Th. 6). This leads to the following translation a conditional $A > B$ in **PCL** into **GL+K**:

$$\Box_K \Box((A \wedge \Box \neg A) \rightarrow B).$$

We will first use this translation as guiding intuition for our calculus, then formally establish soundness and completeness of the calculus independently of these informal

$$\frac{\{\Box((A_i \wedge \Box\neg A_i) \rightarrow B_i)\}_{i=1}^n \vdash \Box((A \wedge \Box\neg A) \rightarrow B)}{\{A_i > B_i\}_{i=1}^n, \Gamma \vdash (A > B), \Delta} (>R) \quad \frac{\Gamma^\Box, \Gamma^{\Box\downarrow}, \Box A \vdash A}{\Gamma \vdash \Box A, \Delta} (\Box R)$$

Figure 3.1: Rules of calculus $SC_{\mathbf{PCL}}$ for logic $\mathbf{PCL}$ (in addition to the $\mathbf{LK}$ rules). Here $\Gamma^\Box = \{\Box A \mid \Box A \in \Gamma\}$ and $\Gamma^{\Box\downarrow} = \{A \mid \Box A \in \Gamma\}$.

considerations, and then will use the calculus to prove that this translation is indeed an embedding of $\mathbf{PCL}$ into $\mathbf{GL+K}$ in the next section.

Our calculus $SC_{\mathbf{PCL}}$ for $\mathbf{PCL}$ extends $\mathbf{LK}$ with the two modal rules shown in Fig. 3.1. The rule $(\Box R)$ is just the (GL) rule of the calculus for $\mathbf{GL}$ (Fig. 2.3, p.20), applied to the $\Box$ -modality. The rule $(>R)$ eliminates a conditional on the right-hand side by applying the conjectured translation to all conditionals in the sequent and then applying the elimination rule $(\Box_K R)$ from the calculus for $\mathbf{K}$ to the resulting $\Box_K$ -modalities. Notice that in $(>R)$ the premise contains only (some of) the translated conditionals from the conclusion. As a consequence, $\Box$ -modalities that may occur in the conclusion are not propagated to the premise and therefore do not mix with the newly introduced $\Box$ -modalities. This separation allows us to safely use a single $\Box$ -modality to represent different preference relations.

Example 3. *The principle of cautious monotonicity, valid in $\mathbf{PCL}$, is given by the following sequent:*

$$(A > B), (A > C) \vdash ((A \wedge B) > C).$$

It can be derived in $SC_{\mathbf{PCL}}$ as follows (for more concise presentation, we only show formulas on the LHS that are relevant for the remaining derivation and do not unfold the subderivation of the top-most classically valid sequent):

$$\frac{\begin{array}{c} \dots, A \vdash C, A \quad (id) \quad \frac{\dots, (A \wedge \Box\neg A) \rightarrow B, \neg(A \wedge B), \Box\neg A \Rightarrow \neg A \quad (\neg R)}{\dots, \Box((A \wedge \Box\neg A) \rightarrow B), A, B, \Box\neg(A \wedge B) \Rightarrow C, \Box\neg A \quad (\Box R)} \\ \frac{\dots, \Box((A \wedge \Box\neg A) \rightarrow B), A, B, \Box\neg(A \wedge B) \Rightarrow C, A \wedge \Box\neg A \quad (\wedge R) \quad \dots, C \vdash C \quad (id)}{\dots, \Box((A \wedge \Box\neg A) \rightarrow B), (A \wedge \Box\neg A) \rightarrow C, A, B, \Box\neg(A \wedge B) \Rightarrow C \quad (\rightarrow L)} \end{array}}{\begin{array}{c} \dots, \Box((A \wedge \Box\neg A) \rightarrow B), (A \wedge \Box\neg A) \rightarrow C, A, B, \Box\neg(A \wedge B) \Rightarrow C \quad (\wedge L) \times 2 \\ \dots, \Box((A \wedge \Box\neg A) \rightarrow B), (A \wedge \Box\neg A) \rightarrow C, (A \wedge B \wedge \Box\neg(A \wedge B)) \Rightarrow C \quad (\rightarrow R) \\ \dots, \Box((A \wedge \Box\neg A) \rightarrow B), (A \wedge \Box\neg A) \rightarrow C \Rightarrow (A \wedge B \wedge \Box\neg(A \wedge B)) \rightarrow C \quad (\Box R) \\ \Box((A \wedge \Box\neg A) \rightarrow B), \Box((A \wedge \Box\neg A) \rightarrow C) \vdash \Box((A \wedge B \wedge \Box\neg(A \wedge B)) \rightarrow C) \quad (\Box R) \\ \hline (A > B), (A > C) \vdash ((A \wedge B) > C) \quad (>R) \end{array}}$$

Before proving soundness and completeness of $SC_{\mathbf{PCL}}$ let us observe some of its properties.

First, $SC_{\mathbf{PCL}}$ does not strictly satisfy the subformula property. Nevertheless, the only formulas introduced that are not subformulas of the root sequent arise from applying the GL-based translation to conditionals occurring in it. Thus, the formulas that may occur in a derivation are still restricted to a finite set determined by the root sequent. We call this finitary restriction the *relaxed subformula property*.

Lemma 3 (Relaxed subformula property). *Every formula occurring in the derivation in $SC_{\mathbf{PCL}}$ with a sequent $\Gamma \vdash \Delta$ in the root is either a subformula of a formula in $\Gamma \cup \Delta$, or a subformula of $\Box((A \wedge \Box \neg A) \rightarrow B)$ for some conditional $A > B$ occurring in a formula in $\Gamma \cup \Delta$.*

Proof. The claim follows immediately by induction on the derivation: the rules of **LK** and $(\Box R)$ introduce only subformulas of formulas in their conclusions, while $(> R)$ additionally introduces only subformulas of $\Box((A \wedge \Box \neg A) \rightarrow B)$ for conditionals $A > B$ occurring in its conclusion. $\square$

The relaxed subformula property, together with the specific formulation of the $(\Box R)$ rule, enables to define a notion of complexity for sequents that strictly decreases during bottom-up derivation.

Definition 3. *The complexity of a sequent $\Gamma \Rightarrow \Delta$ is a triple (d, b, c) , where d is the maximal depth of nesting of conditionals in the sequent, b is a number of $\Box$ -formulas that appear as subformulas in $\Gamma \cup \Delta$ but do not appear as formulas in Γ , and c the total number of propositional connectives in the sequent.*

Using a lexicographic ordering on such triples, the complexity of the premises is strictly smaller than the complexity of the conclusion for any rule application (apart from a redundant application of $(\Box R)$, which contain $\Box A$ on both sides in the conclusion, and therefore can be closed immediately with the concluding rule of **LK**). Specifically, in $(> R)$ parameter d decreases, in $(\Box R)$ parameter b decreases (if the rule application is not redundant) while parameter d decreases or remains the same, and in every rule of **LK** parameter c decreases while both d and b either decrease or remain the same. Since all three parameters are polynomially bounded in the size of the root sequent, the proof search in $SC_{\mathbf{PCL}}$ can be performed with optimal PSPACE complexity. Thus proof search in $SC_{\mathbf{PCL}}$ attempting all non-redundant rule applications in any arbitrary order can be performed in polynomial space, matching the complexity of **PCL**.

Proposition 1. *Proof search in $SC_{\mathbf{PCL}}$ can be performed in PSPACE.*

Proof. By the relaxed subformula property, the size of every sequent is polynomially bounded by the size of the root sequent. Moreover, the complexity from Def. 3 strictly decreases with every non-redundant bottom-up rule application and is polynomially bounded; hence every branch has polynomial length and proof search can be performed by depth-first search using polynomial space. $\square$

We now establish the soundness and completeness of $SC_{\mathbf{PCL}}$ for **PCL**. We do it semantically, by showing that a sequent is derivable in $SC_{\mathbf{PCL}}$ if and only if it is valid in all finite preferential models. To use validity for sequents that may contain $\Box$ -modalities, we need to extend the definition of satisfaction in preferential models to the language

with $\boxminus$ -modalities. Since in our calculus the $\boxminus$ -modality may correspond to different preference relations at different stages of a derivation, we define satisfaction w.r.t. two worlds: a world of evaluation w and a point-of-view world (POV) v , whose preference relation is currently used.

Definition 4. A twice-pointed preferential model is a tuple (M, v, w) where $M = \langle W, R, \{\succeq_w\}_{w \in W}, V \rangle$ is a preferential model, $v \in W$ (called the POV world) and $w \in R(v)$ (called the world of evaluation). Satisfaction of a formula A in the language of **PCL** extended with $\boxminus$ -modalities w.r.t. twice-pointed preferential model (M, v, w) is defined inductively as follows:

- $(M, v, w) \models x$ if $x \in V(w)$;
- $(M, v, w) \models \neg A$ if $(M, v, w) \not\models A$;
- $(M, v, w) \models A \wedge B$ if $(M, v, w) \models A$ and $(M, v, w) \models B$;
- $(M, v, w) \models A \vee B$ if $(M, v, w) \models A$ or $(M, v, w) \models B$;
- $(M, v, w) \models A \rightarrow B$ if $(M, v, w) \not\models A$ or $(M, v, w) \models B$;
- $(M, v, w) \models A > B$ if $(M, w, u) \models B$ for every $u \in \max_{\succ_w}(\|A\|_w^M)$
where $\|A\|_w^M = \{w' \in R(w) : (M, w, w') \models A\}$;
- $(M, v, w) \models \boxminus A$ if $(M, v, u) \models A$ for any $u \in R(v)$ s.t. $w \succ_v u$.

Notice that the POV changes only when evaluating the conditional. Notice also that we use the simpler truth condition for conditionals in terms of most preferred worlds, which is justified for finite preferential models (Lem. 2, p.32). We now define the validity of sequents over the language with $\boxminus$ -modalities w.r.t. finite twice-pointed preferential models.

Definition 5. A sequent $\Gamma \vdash \Delta$ is said to be valid w.r.t. finite twice-pointed preferential models when $(M, v, w) \models (\bigwedge_{A \in \Gamma} A) \rightarrow (\bigvee_{B \in \Delta} B)$ for every twice-pointed preferential model (M, v, w) with a finite set of worlds.

Note that for formulas without $\boxminus$ -modalities, twice-pointed validity coincides with the standard notion of validity.

Lemma 4. If $\Gamma \vdash \Delta$ that does not contain $\boxminus$ -modalities, $\Gamma \vdash \Delta$ is valid w.r.t. twice-pointed finite preferential models iff $M, w \models ((\bigwedge_{A \in \Gamma} A) \rightarrow (\bigwedge_{B \in \Delta} B))$ for every finite preferential model M and any world w in it.

Proof. Let us prove the statement contrapositively. Let $X = ((\bigwedge_{A \in \Gamma} A) \rightarrow (\bigvee_{B \in \Delta} B))$. If $(M, v, w) \not\models X$ for some finite preferential model M , then $M, w \not\models X$, since for the formulas without $\boxminus$ -modalities two-pointed evaluation does not depend on the POV world.

Conversely, if $M, w_0 \not\models X$ for some finite preferential model $M = \langle W, R, \{\succeq_w\}_{w \in W}, V \rangle$ and $w_0 \in M$, consider the following model $M' = \langle W', R', \{\succeq'_w\}_{w \in W'}, V' \rangle$ extending M with an additional world w_1 :

- $W' = W \sqcup \{w_1\}$
- $R'(w_1) = \{w_0\}$;
- $R'(w) = R(w)$ for every $w \in W$;
- $\succeq'_{w_1}$ is a trivial partial order on a singleton set $\{w_0\}$;
- $\succeq'_w = \succeq_w$ for every $w \in W$;
- $V'(w_1) = \emptyset$
- $V'(w) = V(w)$ for every $w \in W$.

Using w_1 as the POV world, it is easy to check that $M, w_0 \not\models X$ implies $(M', w_1, w_0) \not\models X$, since the evaluation of X does not depend on the POV world and is not affected by the additional world w_1 , which is not accessible from any world in W . $\square$

Together with the finite model property for **PCL**, this implies that for any formula A in the language of **PCL**, A is a theorem of **PCL** exactly when sequent $\vdash A$ is valid w.r.t. finite twice-pointed preferential models, allowing us to legitimately use this notion of validity for soundness and completeness. We begin with soundness.

Theorem 6 (Soundness of $SC_{\mathbf{PCL}}$). *If $\Gamma \vdash \Delta$ is derivable in $SC_{\mathbf{PCL}}$ then $\Gamma \vdash \Delta$ is valid w.r.t. twice-pointed finite preferential models.*

Proof. The proof is by induction on the derivation, for each rule we show that the validity of all premises implies the validity of the conclusion. For rules of **LK** this is trivial.

For rule $(\boxminus R)$: Suppose towards a contradiction that the conclusion is not valid. Then there is a finite twice-pointed model (M, v, w) with $M = \langle W, R, \{\succeq_w\}_{w \in W}, V \rangle$ s.t. all the formulas in Γ are true at (M, v, w) while all formulas in Δ and $\boxminus A$ are false at (M, v, w) . The latter implies that the set $U = \{u \in R(v) \mid u \succ_v w \text{ and } (M, u, v) \not\models A\}$ is not empty. Then there is some $u \in \max_{\succ_v}(U)$, since $\succ_u$ is a strict preorder on a finite set. (M, v, u) satisfies all formulas in $\Gamma^{\boxminus \downarrow}$ and $\Gamma^{\boxminus}$ due to the transitivity of $\succ_u$, it also falsifies A by definition of U and satisfies $\boxminus A$ since $u \in \max_{\succ_v}(U)$. Thus, (M, v, u) falsifies the premise of $(\boxminus R)$, leading to a contradiction.

For rule $(> R)$: Suppose towards a contradiction that the conclusion is not valid. Then there is a finite twice-pointed model (M, v_0, w_0) with $M = \langle W, R, \{\succeq_w\}_{w \in W}, V \rangle$ s.t. (M, v_0, w_0) satisfies $(A_i > B_i)$ for every $1 \leq i \leq n$ and falsifies $(A > B)$. The latter implies that there is $w_1 \in \max_{\succ_{w_0}}(\|A\|_{w_0}^M)$ s.t. $w_1 \notin \|B\|_{w_0}^M$. Let us consider the following

model $M' = \langle W', R', \{\succeq'_w\}_{w \in W'}, V' \rangle$ that adds a new "dummy" world w'_0 to M , which takes the set of accessible worlds and the preference relation from w_0 and modifies them by making w'_0 itself also accessible from w'_0 and the least preferred among the accessible worlds. Namely,

- $W' = W \sqcup \{w'_0\}$
- $R'(w'_0) = R(w_0) \cup \{w'_0\}$
- $R'(w) = R(w)$ for any $w \in W$
- $x \succeq_{w'_0} y$ iff either $x \succeq_{w_0} y$ or $y = w'_0$
- $\succeq'_w = \succeq_w$ for every $w \in W$
- $V'(w'_0) = \emptyset$
- $V'(w) = V(w)$ for every $w \in W$

It is easy to see by straightforward induction that $(M', v, w) \models C$ is equivalent to $(M, v, w) \models C$ for any $v, w \in W$ and any formula C , since the additional world w'_0 is not accessible from the original worlds in W and does not affect the evaluation. An analogous straightforward inductive argument further shows that $(M', w'_0, w) \models C$ is equivalent to $(M, w_0, w) \models C$ for any $w \in R(w_0)$ and any formula C , since $\succeq_{w'_0}$ and $\succeq_{w_0}$ coincide on $R(w_0)$. Using this fact, we can show that the twice-pointed model (M', w'_0, w'_0) falsifies the premise, as required for a desired contradiction. First, for any $1 \leq i \leq n$, take any world $u \in R'(w'_0)$ s.t. $u \succ_{w'_0} w'_0$ (thus, $u \in R(w_0)$). If both $(M', w'_0, u) \models A_i$ and $(M', w'_0, u) \not\models \Box \neg A_i$ then by the equivalence above we get that $(M, w_0, u) \models A_i$ and $(M, w_0, u) \not\models \Box \neg A_i$, implying that $u \in \max_{\succ_{w_0}} (\|A_i\|_{w_0}^M)$, and since we assumed that $(M, v_0, w_0) \models A_i > B_i$, we get that $(M, w_0, u) \models B_i$, which can be rewritten back by the same equivalence as $(M', w'_0, u) \models B_i$. This argument shows that $(M', w'_0, w_0) \models \Box((A_i \wedge \Box \neg A_i) \rightarrow B_i)$ for every $1 \leq i \leq n$. At the same time, assumption $(M, v_0, w_0) \not\models A > B$, implies that there is a world $u_0 \in R(w_0)$ s.t. $(M, w_0, u_0) \models A \wedge \Box \neg A$ and $(M, w_0, u_0) \not\models B$, which by the same equivalence gives us $(M', w'_0, u_0) \models A \wedge \Box \neg A$ and $(M', w'_0, u_0) \not\models B$, implying $(M', w'_0, w'_0) \not\models \Box((A \wedge \Box \neg A) \rightarrow B)$ since $u_0 \succ_{w'_0} w'_0$ by the definition of $\succ_{w'_0}$. $\square$

Now we show completeness of $SC_{\mathbf{PCL}}$, by constructing a countermodel for any non-derivable sequent.

Theorem 7 (Completeness of $SC_{\mathbf{PCL}}$). *If $\Gamma \vdash \Delta$ is valid w.r.t. finite twice-pointed preferential models then $\Gamma \vdash \Delta$ is derivable in $SC_{\mathbf{PCL}}$.*

Proof. We prove the statement contrapositively: assuming that $\Gamma \vdash \Delta$ is not derivable in $SC_{\mathbf{PCL}}$, we construct a finite twice-pointed model falsifying it. The proof proceeds by induction on the complexity of $\Gamma \Rightarrow \Delta$.

If there is a formula in $\Gamma \cup \Delta$ that is neither a propositional atom nor a conditional nor a $\boxminus$ -formula, then an elimination rule of **LK** can be applied to it. At least one of the premises is a non-derivable sequent of smaller complexity, so by the inductive hypothesis there will be a countermodel falsifying it, which will also falsify the conclusion.

Now let all formulas in $\Gamma \cup \Delta$ be either a propositional atom, a conditional, or a $\boxminus$ -formula. The premise of each application of $(>R)$ or $(\boxminus R)$ is an underivable sequent of smaller complexity, so by inductive hypothesis, there exists a twice-pointed model falsifying it. Suppose $(M^{+1}, v^{+1}, w^{+1}), \dots, (M^{+n}, v^{+n}, w^{+n})$ are all such models for all possible applications of $(\boxminus R)$, and $(M^{-1}, v^{-1}, w^{-1}), \dots, (M^{-m}, v^{-m}, w^{-m})$ are all such models for all possible applications of $(>R)$. Denote $I = \{+1, \dots, +n, -1, \dots, -m\}$. Let $M^i = \langle W^i, R^i, \{\succeq_w^i\}_{w \in W^i}, V^i \rangle$ for $i \in I$. We will construct our countermodel for $\Gamma \Rightarrow \Delta$ as the disjoint union of all these countermodels, adding an external world of evaluation w_0 and an external POV world v_0 to them. We will not modify the sets of accessible worlds $R^i(w)$ and orderings $\succeq_w^i$ for any world of any submodel M^i , and define it separately only for the additional worlds w_0 and v_0 . Specifically, we consider the model $M = \langle W, R, \{\succeq_w\}_{w \in W}, V \rangle$, defined as follows:

- $W = \{w_0, v_0\} \sqcup W^{+1} \sqcup \dots \sqcup W^{+n} \sqcup W^{-1} \sqcup \dots \sqcup W^{-m}$;
- $R(w_0) = \bigcup_{1 \leq i \leq m} \{w \in R^{-i}(v^{-i}) : w \succeq_{v^{-i}} w^{-i}\}$;
- $R(v_0) = W$;
- $R(w) = R^k(w)$ for any $k \in I$ and $w \in W^k$;
- $x \succeq_{w_0} y$ iff there is $i \in \{1, \dots, m\}$ s.t. $x, y \in R^{-i}(v^{-i}) \cap R(w_0)$ and $x \succeq_{v^{-i}} y$
- and $x \succeq_{v_0} y$ in the following four cases:
 1. $x, y \in W^k$ for some $k \in I$ and $x \succeq_{v^k} y$;
 2. $x = y = v_0$;
 3. $x = y = w_0$;
 4. $y = w_0, x \in W^{+i}$, and $x \succeq_{v^{+i}} w^{+i}$;
- $x \succeq_w y$ iff $x \succeq_w^k y$ for any $k \in I$ and $w \in W^k$;
- $V(w_0) = \Gamma \cap \text{Atm}$;
- $V(v_0) = \emptyset$;
- $V(w) = V^k(w)$ for $k \in I$ and $w \in W^k$.

First, notice that W is finite and all relations $\succeq_w$ are reflexive, antisymmetric, and transitive (if the same is assumed for all submodels M^k). Also, it is straightforward to show that $(M, u, w) \models C$ iff $(M^k, u, w) \models C$ for any formula C and any $k \in I$

and $u, w \in W^k$ (since relations and valuation within submodels did not change and the evaluation of the formula in a submodel does not access the worlds outside this submodel). Moreover, $(M, v_0, w) \models C$ iff $(M^{+i}, v^{+i}, w) \models C$ for any $i \in \{1, \dots, n\}$ and $w \in R^{+i}(v^{+i})$ (since $\succeq_{v_0}$ and $\succeq_{v^{+i}}$ coincide on W^{+i}), and $(M, w_0, w) \models C$ iff $(M^{-i}, v^{-i}, w) \models C$ for any $i \in \{1, \dots, m\}$ and $w \in R(w_0) \cap R^{-i}(v^{-i})$ (since $\succeq_{w_0}$ and $\succeq_{v^{-i}}$ coincide on $R(w_0) \cap W^{-i}$). Relying on these equivalences, we can show that (M, v_0, w_0) satisfies (resp. falsifies) all formulas in Γ (resp. Δ):

- for propositional atoms, due to the definition of V and the fact that $\Gamma \vdash \Delta$ can not be closed by the concluding rule of **LK**;
- for $\Box A$ on the LHS due to presence of both A and $\Box A$ on the LHS in the premise of any application of $(\Box R)$, which implies $(M^{+i}, v^{+i}, w^{+i}) \models A$ and $(M^{+i}, v^{+i}, w) \models A$ for all $w \in R^{+i}(v^{+i})$ s.t. $w \succ_{v^{+i}} w^{+i}$ (which covers all worlds $w \in W$ s.t. $w \succ_{v_0} w_0$ due to antisymmetry);
- for $\Box A$ on the RHS due to possible application of $(\Box R)$ to this formula with A on the right in this premise, implying that $(M, v^{+i}, w^{+i}) \not\models A$ for some $i \in \{1, \dots, n\}$, so $(M, v_0, w^{+i}) \not\models A$ by equivalence above, so $(M, v_0, w_0) \not\models \Box A$ since $w^{+i} \succ_{v_0} w_0$;
- for $A > B$ on the LHS due to the presence of $\Box(A \wedge \Box \neg A \rightarrow B)$ on the LHS of the premise of any application of $(> R)$, which ensures that for every $i \in \{1, \dots, m\}$,

$$\max_{\succ_{w_0}}(|A|_{w_0}^M) \cap R^{-i}(v^{-i}) = \max_{\succ_{v^{-i}}}(|A|_{v^{-i}}^{M^{-i}}) \cap R(w_0) \subseteq |B|_{v^{-i}}^{M^{-i}} \subseteq |B|_{w_0}^M,$$

which implies $\max_{\succ_{w_0}}(|A|_{w_0}^M) \subseteq |B|_{w_0}^M$ since $R(w_0) \subseteq \bigcup_i R^{-i}(v^{-i})$.

- for $(A > B)$ on the RHS due to a possible application of $(> R)$ to this formula with $\Box(A \wedge \Box \neg A \rightarrow B)$ on the RHS in this premise, implying that there is $z \in R^{-i}(v^{-i})$ s.t. $z \succ_{v^{-i}} w^{-i}$ and $(M^{-i}, v^{-i}, z) \models A \wedge \Box \neg A$ and $(M^{-i}, v^{-i}, z) \not\models B$. By one of the equivalences above this implies $z \in \max_{\succ_{w_0}}(|A|_{w_0}^M)$ and $z \notin |B|_{w_0}^M$, thus $(M, v_0, w_0) \not\models A > B$. $\square$

Remark 1. *The countermodel constructed in the above proof has the same structure as the Friedman-Halpern small model construction for **PCL** [FH94]: starting with a formula that has no $\Box$ -modalities, the resulting model will consist of separate chains of worlds, each chain falsifies one conditional in some world of a later chain. While Friedman-Halpern small models are obtained non-constructively via selecting a finite subset of worlds from an arbitrary existing model, our calculus provides a constructive way of obtaining such small models by analyzing failed derivations.*

3.5 Bridge: Embedding of PCL into GL+K

Having established the analytic calculus for **PCL**, we can now formally state and prove an embedding of **PCL** into the bimodal logic **GL+K** that guided its development.

Bimodal logics combine two modalities governed by different logical systems. Accordingly, the language of **GL+K** extends propositional logic with two modal operators $\Box_{\mathbf{GL}}$ and $\Box_{\mathbf{K}}$:

$$\mathcal{L}_{\mathbf{GL+K}} = A ::= x \in \text{Atm} \mid \neg A \mid A \wedge A \mid A \vee A \mid A \rightarrow A \mid \Box_{\mathbf{GL}} A \mid \Box_{\mathbf{K}} A.$$

Bimodal logics are defined semantically via relational models equipped with two independent accessibility relations, each satisfying the frame conditions of the corresponding logic. Each modal operator is then evaluated using its respective accessibility relation.

Definition 6. A **GL+K**-model is a tuple $M = \langle W, R_{\mathbf{K}}, R_{\mathbf{GL}}, V \rangle$, where W is a finite set of worlds, $R_{\mathbf{K}}$ is arbitrary binary relation on W , $R_{\mathbf{GL}}$ is transitive and asymmetric binary relation on W , and V is valuation on W . Satisfaction of formulas from $\mathcal{L}_{\mathbf{GL+K}}$ w.r.t. **GL+K**-models is defined as follows:

$$\begin{array}{lll} M, w \models x & \text{iff} & x \in V(w) \\ M, w \models \neg A & \text{iff} & M, w \not\models A \\ M, w \models A \wedge B & \text{iff} & M, w \models A \text{ and } M, w \models B \\ M, w \models A \vee B & \text{iff} & M, w \models A \text{ or } M, w \models B \\ M, w \models A \rightarrow B & \text{iff} & M, w \not\models A \text{ or } M, w \models B \\ M, w \models \Box_{\mathbf{K}} A & \text{iff} & M, u \models A \text{ for any } u \in W \text{ s.t. } w R_{\mathbf{K}} u \\ M, w \models \Box_{\mathbf{GL}} A & \text{iff} & M, u \models A \text{ for any } u \in W \text{ s.t. } w R_{\mathbf{GL}} u \end{array}$$

The embedding of **PCL** into **GL+K** is defined by the following mapping $tr_{\mathbf{PCL}}$ from $\mathcal{L}_{\mathbf{PCL}}$ to $\mathcal{L}_{\mathbf{GL+K}}$:

$$\begin{array}{ll} tr_{\mathbf{PCL}}(p) & = p \\ tr_{\mathbf{PCL}}(\neg A) & = \neg tr_{\mathbf{PCL}}(A) \\ tr_{\mathbf{PCL}}(A \wedge B) & = tr_{\mathbf{PCL}}(A) \wedge tr_{\mathbf{PCL}}(B) \\ tr_{\mathbf{PCL}}(A \vee B) & = tr_{\mathbf{PCL}}(A) \vee tr_{\mathbf{PCL}}(B) \\ tr_{\mathbf{PCL}}(A \rightarrow B) & = tr_{\mathbf{PCL}}(A) \rightarrow tr_{\mathbf{PCL}}(B) \\ tr_{\mathbf{PCL}}(A > B) & = \Box_{\mathbf{K}} \Box_{\mathbf{GL}} ((tr_{\mathbf{PCL}}(A) \wedge \Box_{\mathbf{GL}} \neg tr_{\mathbf{PCL}}(A)) \rightarrow tr_{\mathbf{PCL}}(B)) \end{array}$$

We can prove that this translation constitutes a faithful embedding using the correspondence between the rules ($\boxtimes R$) and ($> R$) and modal rules for elimination of $\Box_{\mathbf{K}}$ and $\Box_{\mathbf{GL}}$.

Theorem 8. A is valid in **PCL** iff $tr_{\mathbf{PCL}}(A)$ is valid in **GL+K**.

Proof. ($\Leftarrow$): if A valid in **PCL**, by Th. 7 there exists a derivation for $\vdash A$ in $SC_{\mathbf{PCL}}$, and we can easily show that if we apply translation $tr_{\mathbf{PCL}}$ to every formula in this derivation, each rule application will preserve validity w.r.t. **GL+K**-models (since the rule ($> R$) after translation of conditionals in the conclusion behaves exactly like the standard sequent rule for $\Box_{\mathbf{K}}$ -modality).

($\Rightarrow$): if $tr_{\mathbf{PCL}}(X)$ is valid in $\mathbf{GL+K}$, it can be derived using the rules of $\mathbf{LK}$ plus the rule $(\Box_K R)$ for $\Box_K$ and (GL) for $\Box_{GL}$. This derivation can be transformed into a derivation of X in $SC_{\mathbf{PCL}}$. *For this, replace all $\Box_{GL}$ -modalities with $\Box$ -modalities.* This allows to replace each application of (GL) with identical application of $(\Box R)$. As for applications of $(\Box_K R)$ notice that all $\Box_K$ -modalities in $tr_{\mathbf{PCL}}(X)$ have formula of the form $\Box_{GL}((A \wedge \Box_{GL} \neg A) \rightarrow B)$ under it, so the rule application

$$\frac{\{\Box_{GL}((A_i \wedge \Box_{GL} \neg A_i) \rightarrow B_i)\}_{i=1}^n \vdash \Box_{GL}((A \wedge \Box_{GL} \neg A) \rightarrow B)}{\{\Box_K \Box_{GL}((A \wedge \Box_{GL} \neg A) \rightarrow B)\}_{i=1}^n, \Gamma \vdash \Box_K \Box_{GL}((A \wedge \Box_{GL} \neg A) \rightarrow B), \Delta} (\Box_K R)$$

can always be replaced with

$$\frac{\{\Box((A_i \wedge \Box \neg A_i) \rightarrow B_i)\}_{i=1}^n \vdash \Box((A \wedge \Box \neg A) \rightarrow B)}{\{A_i > B_i\}_{i=1}^n, \Gamma \vdash (A > B), \Delta} (> R)$$

This transformation gives us exactly the derivation of X in $SC_{\mathbf{PCL}}$, which implies that it is valid in $\mathbf{PCL}$. $\square$

As the KLM logic $\mathbf{P}$ coincides with the flat fragment of $\mathbf{PCL}$ the adequacy of $tr_{\mathbf{PCL}}$ also justifies the simplification of the embedding of $\mathbf{P}$ into $\mathbf{GL}$ in [GGOP09]: the entailment

$$\{A_i > B_i\}_i \vdash A > B$$

holds in $\mathbf{P}$ (and, so, in $\mathbf{PCL}$) iff

$$\{\Box_K \Box(A_i \wedge \Box \neg A_i \rightarrow B_i)\}_i \vdash (\Box_K \Box(A \wedge \Box \neg A \rightarrow B))$$

holds in $\mathbf{GL+K}$, which is equivalent in $\mathbf{GL+K}$ to

$$\{\Box(A_i \wedge \Box \neg A_i \rightarrow B_i)\}_i \vdash (\Box(A \wedge \Box \neg A \rightarrow B)),$$

which in turn is equivalent to the same entailment in $\mathbf{GL}$.

Part II

Norms

Preference-Based Deontic Logics

We now turn to another interpretation of conditionals: the deontic one. Deontic logic is concerned with the formal study of normative concepts such as obligation, permission, and prohibition. Originally developed to clarify the structure of moral and legal reasoning in philosophy, it now finds important applications as a formal framework for automated normative reasoning — for instance, in norm-compliance verification for multi-agent systems, in normative AI, or in the development of legal decision-support tools [BvdTV07].

Traditionally, deontic concepts were modelled using the tools of standard modal logic, culminating in the Standard Deontic Logic (**SDL**) [vW51], where obligation is taken as the primary concept and modelled as a unary modality of the normal modal logic **D**. **SDL**, however, proved too rigid for real-world normative reasoning, failing to handle complex scenarios involving conflicting or context-dependent norms. One well-known difficulty is the *contrary-to-duty* problem [Chi63], where the failure of complying with an obligation (primary obligation), trigger another obligation (secondary obligation). **SDL** cannot simultaneously account for the primary and the secondary, contrary-to-duty obligation. Similar issues arise with handling *specificity* of norm application, when a general obligation should be overridden by a more specific one. These limitations were a major driver of subsequent developments in the field, and it was gradually recognized that representing obligations as conditionals $\bigcirc(B|A)$ —read as “ B is obligatory under the condition A ”— provides a natural setting for resolving them. The conditional approach has subsequently become the prevalent paradigm in deontic logic [GHP⁺13].

The literature offers a big variety of formalisms for conditional obligation, which can be broadly divided into two categories [Par26]: *preference-based* and *norm-based* approaches. Preference-based systems employ preferential semantics to express obligation in terms of degrees of “ideality” of worlds, providing a natural framework for handling norm violation and contrary-to-duty obligations. Norm-based systems, by contrast, are operational in character, defining obligations relative to an explicit set of norms, and are particularly well-suited to handle specificity. The two approaches are not only conceptually distinct but

also technically quite remote, relying on different semantic primitives and mathematical tools, which makes direct comparison and unification rather difficult. In this and the following chapter, we examine both from an analytic proof-theoretic perspective, offering an abstract and constructive treatment of prominent formalisms within each tradition — Åqvist’s logics and Input/Output logics.

In this chapter, we begin with the preference-based approach, which served as the foundational method for introducing deontic conditional logic. We develop analytic characterizations for Åqvist’s systems, the most representative family of logics in this approach. After surveying the historical background (Section 4.1), the state of the art regarding analytic methods (Section 4.2), and formally introducing Åqvist’s logics (Section 4.3), we develop new small-model constructions (Section 4.4)—*LEGO models*—which are suitable for relaxed versions of preference relation and cover uniformly the whole Åqvist’s family. In Sec. 4.5, we use LEGO models to obtain new semantical characterizations for Åqvist’s via classes of finite models defined by simple frame conditions, resolving the conjecture from [COP⁺23] that **F** can be characterized by finite models with acyclic preference relation. In Sec. 4.6, we deduce NP-completeness for all Åqvist’s logics from our small-model theorems. These results serve as the basis for the automated prover DeoSMT¹, described in Sec. 4.7. Relying on the new semantical characterization, in Sec. 4.8 we extend the **GL**-based methodology of [GGOP09] and the previous chapter to the hypersequent framework, obtaining a uniform proof-theoretic treatment of all Åqvist systems. Finally, in Sec. 4.9 we present new embeddings of **F**, **F_{CM}** and **G** into normal bimodal logics, extending the existing embedding for **E** from [COP22].

4.1 Background

Originally proposed by Hansson [Han69] to address the complexities of contrary-to-duty obligations, the preference-based approach draws heavily on rational choice theory. In parallel to Lewis and Burgess, Hansson introduced preferential models, where the preference relation is interpreted as the “comparative goodness” of worlds, and proposed to define the truth of conditional obligation $O(B|A)$ as the fact that B holds in all “best” worlds among those satisfying A . This approach was significantly refined by Åqvist [Åqv84], who studied and axiomatized various versions of the preference relation within a deontic context, while also combining the deontic modality $\bigcirc(B|A)$ with the alethic modality $\Box A$ to integrate normative and factual reasoning within a single framework. This resulted in a hierarchical family of logics with increasing deductive strength: system **E** imposes no conditions on the preference relation (not even assuming it is an order), system **F** restricts the relation minimally to rule out contradictory obligations via the so-called *limitedness* property, and system **G** assumes the relation to be a total order, mirroring Hansson’s original version. The family was later extended [Par14] by inserting between **F** and **G** system **F_{CM}**, where the total orders of **G** are generalized

¹The prover has been developed by Christian Köll, a master student co-supervised by A. Ciabattoni and myself

to partial orders, similarly to Burgess’s generalization of Lewis’s counterfactual models. Axiomatically, this is captured by extending the system **F** with the well-known cautious monotonicity axiom [Gab85], hence the name **F_{CM}**. On a technical level, the higher end of the Åqvist’s spectrum intersects with logics studied in counterfactual contexts: **G** and **F_{CM}** can be viewed as deontic interpretation of the counterfactual logics **VTA** and **PCLTA**, respectively. The latter two extend Lewis’s logic **V** [Lew73] and Burgess’s logic **PCL** [Bur81] with frame conditions of reflexivity and absoluteness, resulting in a single preference relation that covers all worlds in the model. At the same time, the relaxed notion of a non-order preference relation in **E** and **F** remains somewhat unique to the deontic context: while transitivity is naturally assumed for “closeness” or “normality” interpretations in other conditional logics, there are significant conceptual arguments against the transitivity of “comparative goodness” [HGY22, PS14]. Over time, Åqvist’s family has become the de-facto standard in preference-based deontic logic, providing a robust scale against which various deontic scenarios can be evaluated.

4.2 State of the art

Åqvist’s systems have been extensively studied from semantic and axiomatic perspectives, resulting in a variety of alternative characterizations based on different conditions on preference models, axiom systems, and notions of “bestness”, surveyed in [Par21]. More recently, attention has shifted toward the computational perspective, motivated by potential applications in automated decision-making. Decidability of theoremhood for all four logics was established [Par21] via an alternative semantics based on selection functions, and an encoding [BFMP19] into the highly expressive (and potentially undecidable) framework of Higher-Order Logic was proposed as the basis for automated deduction and implemented in Isabelle/HOL. While this provides a powerful environment for experimentation and meta-reasoning, it offers little insight into the intrinsic computational complexity of these logics, and is not well suited for efficient object-level reasoning tasks such as entailment or consistency checking. These issues motivated the quest for analytic characterizations of Åqvist’s systems.

On the semantic side, the logics **VTA** and **PCLTA** (i.e. **G** and **F_{CM}**) are already covered by the small-model constructions of Friedman and Halpern [FH94], yielding co-NP-completeness. However, these constructions rely crucially on transitivity of the preference relation and thus do not apply to the weaker systems **E** and **F**. Instead, analytic characterizations for Åqvist’s family have primarily been pursued on the proof-theoretic side. Standard sequent calculi are not sufficient for this purpose, since all Åqvist’s logics subsume modal logic **S5**, which is known to resist natural cut-free sequent characterizations. Proof calculi for Åqvist’s logics was therefore developed in terms of more expressive structures: labeled calculi [GGOS09, GNO21] and sequent calculi with blocks [GLOP16] have been used to develop analytic proof theory for **VTA** and **PCLTA** (i.e. **G** and **F_{CM}**); adapting an approach from [GGOP09], hypersequent calculi with the additional $\boxminus$ -modality in the language have been developed for **E** [COP22] and **F** [COP⁺23]; a similar approach [CT24] was subsequently applied for **G** and **F_{CM}**

to obtain both hypersequent calculi and sequent calculi with analytic cuts, relying on alternative truth condition. However, these results do not provide a uniform proof-theoretic treatment of all four systems. Moreover, while the hypersequent calculus for **E** shows that it is also co-NP-complete, its extension to **F** relies on a technically involved treatment of the non-normal $\boxtimes$ modality together with a syntactic completeness proof based on a sophisticated cut-elimination argument generalizing Valentini's proof [Val83] for **GL**. This approach neither provides a method for extracting countermodels from failed derivations nor resolves the open question of computational complexity of **F**, establishing only a NEXPTIME upper bound. This highlights the difficulty of handling the distinctive limitedness condition of **F**, both semantically and proof-theoretically. In [COP⁺23], it was therefore suggested that a more tractable proof theory for **F** might arise from some alternative semantic characterization with simpler frame conditions, and the combination of finiteness and acyclicity was proposed as a potential such characterization.

4.3 Formal preliminaries: Åqvist's logics

Language. The syntax of Åqvist's logics extends the usual propositional language with two modalities: unary $\Box$ for necessity and binary $\bigcirc(\cdot|A)$ for conditional obligation.

$$\mathcal{L}_{\mathbb{A}q} = A ::= x \in \text{Atm} \mid \neg A \mid A \wedge A \mid A \vee A \mid A \rightarrow A \mid \Box A \mid \bigcirc(A|A).$$

Other normative concepts can be defined via obligation: *conditional prohibition* is defined as $\bigcirc(\neg B|A)$, and *conditional permission* as $\neg \bigcirc(\neg B|A)$.

We will denote by $|A|$ the size of the formula (number of symbols), denote by $\text{Sub}\mathcal{F}(A)$ the set of all subformulas of A (including A), and $\text{Cond}(A) = \{C \mid \bigcirc(D|C) \in \text{Sub}\mathcal{F}(A)\}$ will denote the set of conditions of all obligations appearing in the formula.

Axiomatic characterization. The four logics can also be presented modularly by Hilbert-style axiomatizations [Par21]. Their common basis is **S5**, extended with the following axioms for conditional obligation:

$$\begin{aligned} (COK) \quad & \bigcirc(B \rightarrow C \mid A) \rightarrow (\bigcirc(B \mid A) \rightarrow \bigcirc(C \mid A)) \\ (Id) \quad & \bigcirc(A \mid A), \\ (Sh) \quad & \bigcirc(C \mid A \wedge B) \rightarrow \bigcirc(B \rightarrow C \mid A), \\ (ONec) \quad & \Box A \rightarrow \bigcirc(A \mid B), \\ (Abs) \quad & \bigcirc(B \mid A) \rightarrow \Box \bigcirc(B \mid A), \\ (Ext) \quad & \Box(A \leftrightarrow B) \rightarrow (\bigcirc(C \mid A) \leftrightarrow \bigcirc(C \mid B)). \end{aligned}$$

These axioms define the basic system **E**. The remaining systems are obtained by adding the following axioms:

$$\begin{aligned} (D*) \quad & \bigcirc(\perp \mid A) \rightarrow \Box \neg A, \\ (CM) \quad & \bigcirc(B \mid A) \wedge \bigcirc(C \mid A) \rightarrow \bigcirc(C \mid A \wedge B), \\ (RM) \quad & \neg \bigcirc(\neg B \mid A) \wedge \bigcirc(C \mid A) \rightarrow \bigcirc(C \mid A \wedge B). \end{aligned}$$

$\mathbf{F}$ is obtained by extending $\mathbf{E}$ with axiom (D^*) . $\mathbf{F}_{\mathbf{CM}}$ and $\mathbf{G}$ further extend $\mathbf{F}$ with (CM) and (RM) respectively, Axiom (RM) entails (CM) in $\mathbf{F}$, so these systems form an increasing hierarchy

$$\mathbf{E} \subseteq \mathbf{F} \subseteq \mathbf{F}_{\mathbf{CM}} \subseteq \mathbf{G}.$$

The axioms common to all four systems capture the basic behaviour of conditional obligation. In particular, (COK) is the conditional analogue of the modal axiom $\mathbf{K}$, $(ONec)$ lifts settled truths to obligations under arbitrary conditions, (Ext) allows necessarily equivalent conditions to be exchanged, and (Abs) reflects the world-independent nature of the deontic preference relation.

The additional axioms distinguishing the systems have natural deontic readings. Axiom (D^*) , distinguishing $\mathbf{F}$ from $\mathbf{E}$, rules out conflicting obligations under possible conditions. Axiom (CM) , distinguishing $\mathbf{F}_{\mathbf{CM}}$ from $\mathbf{F}$, expresses *cautious monotonicity*: if both B and C are obligatory given A , then C remains obligatory given $A \wedge B$. Finally, $(\)$, characteristic of $\mathbf{G}$, expresses the stronger principle of *rational monotonicity*: it suffices that B is permitted given A for C to remain obligatory given $A \wedge B$. Both (CM) and (RM) are standard principles of non-monotonic reasoning, discussed further in Chapters 6 and 7, respectively.

Semantics. Semantics of logics in the Åqvist family is also based on preferential models. There are, however, several differences from the counterfactual preferential models considered in the previous chapter. First, there is a single preference relation defined over the whole set W , rather than world-dependent preference relations defined over accessible worlds.² Second, in general—and in particular for the weakest logic $\mathbf{E}$ —no restrictions are imposed on the preference relation. Third, in the deontic setting, conditionals are evaluated using the simpler truth condition in terms of most preferred worlds (see Lem. 2), while the existence of such worlds is controlled by additional conditions distinguishing the different logics in the family, introduced below. This leads to the following definition of preferential models in the deontic setting.

Definition 7. A (deontic) preferential model is a tuple $\langle W, \succeq, V \rangle$, where W is a set of worlds, $\succeq$ is a binary relation on W , and $V : W \rightarrow 2^{Atm}$ is a valuation function. The strict version of $\succeq$, is defined as follows: $v \succ u$ when $v \succeq u$ and $u \not\succeq v$. Satisfaction of a modal formula A in a given world $w \in W$ of a preferential model $M = \langle W, R, \{\succeq_w\}_{w \in W}, V \rangle$ is

²In the counterfactual setting, such simplified models can be obtained by imposing the frame conditions of reflexivity (corresponding to Lewis's axiom T) and absoluteness (corresponding to Lewis's axiom A), hence the names **PCLTA** and **VTA** for the corresponding counterfactual logics.

Logic	Limit conditions		Properties of $\succeq$	
	limited	smooth	transitive	total
E				
F	✓			
F_{CM}		✓	✓	
G		✓	✓	✓

Figure 4.1: Preference-semantical characterizations for Åqvist’s logics [Par21, Tables 1 and 2] (with maximality as the notion of bestness).

defined inductively as follows:

- $M, w \models x$ if $x \in V(w)$;
- $M, w \models \neg A$ if $M, w \not\models A$;
- $M, w \models A \wedge B$ if $M, w \models A$ and $M, w \models B$;
- $M, w \models A \vee B$ if $M, w \models A$ or $M, w \models B$;
- $M, w \models A \rightarrow B$ if $M, w \not\models A$ or $M, w \models B$;
- $M, w \models \Box A$ if $M, u \models A$ for all $u \in W$;
- $M, w \models \bigcirc(B|A)$ if $M, u \models A$ for all $u \in \max_{\succ}(\|A\|^M)$.

where $\|A\|^M = \{w : M, w \models A\}$ is the truth set of formula A and $\max_{\succ}(U) = \{v \in U \mid \nexists u \in U : u \succ v\}$ is the set of best worlds in U w.r.t. $\succ$. We say that M validates A (denoted $M \models A$) when $\|A\|^M = W$.

Notice that the satisfaction of both $\Box A$ and $\bigcirc(D|C)$ does not depend on the world of evaluation.

Different Åqvist logics are defined by different classes of preferential models, obtained by imposing additional properties on the preference relation $\succeq$. There are various alternative characterizations of Åqvist logics based on different combinations of such properties and different notions of “bestness” (see [Par21] for a survey); here we chose one characterization from [Par21] most suitable for our analytic purposes, presented in Fig. 4.1. It employs two standard properties of relations: transitivity ($\succeq$ is transitive if $w_1 \succeq w_2$ and $w_2 \succeq w_3$ imply $w_1 \succeq w_3$) and totality ($\succeq$ is total if for any $w_1, w_2 \in W$, either $w_1 \succeq w_2$ or $w_2 \succeq w_1$). In addition, it uses two *limit properties* governing the existence of best worlds in the model: *limitedness* and *smoothness*.

Definition 8. (*Limit conditions*) Let $M = (W, \succeq, V) \in \mathcal{M}$. M is *limited* when for any formula A if $\|A\|^M \neq \emptyset$ then $\max_{\succ}(\|A\|^M) \neq \emptyset$. M is *smooth* when for any formula A and any world $w \in \|A\|^M$ there exists $u \in \max_{\succ}(\|A\|^M)$ such that either $u = w$ or $u \succ w$.

Note that both limit conditions are not *frame conditions*, in the sense that they depend not only on the preference relation, but also on the valuation in the model.

The semantical characterization of Åqvist's logics is provided in Fig. 4.1. The semantic conditions can be viewed as increasingly strong assumptions on the underlying preference ordering. *Limitedness*, characteristic of $\mathbf{F}$, requires every non-empty set of A -worlds to contain at least one best world, preventing obligations from becoming vacuously true when no optimal A -world exists. *Smoothness*, used for $\mathbf{F}_{\mathbf{CM}}$ and $\mathbf{G}$, strengthens this by requiring every non-best A -world to be dominated by some best one, so that the best worlds genuinely represent the normative ideal among the A -alternatives. Together with transitivity, this supports the stability of obligations under the cautious strengthening expressed by (CM) ; adding totality yields the stronger stability captured by rational monotonicity (RM) in $\mathbf{G}$.

Definition 9. *Formula A is a theorem of Åqvist's logic L iff $M \models A$ for any preference model M that satisfies model conditions for logic L in Fig. 4.1.*

We will call a preference model M a *countermodel* for a formula A if $M \not\models A$, and we will further call it an *L -countermodel* if it belongs to a class of models corresponding to a logic L from Fig. 4.1.

4.4 LEGO model constructions for Åqvist's logics

In this section, we provide a semantic analytic characterization of Åqvist logics via *LEGO* small-model constructions. More specifically, for every logic L from the Åqvist family, we show how an arbitrary L -countermodel M for a formula A can be transformed into a L -countermodel whose number of worlds is bounded polynomially in $|A|$. This transformation is obtained by assembling a small countermodel from elementary building blocks-submodels, using worlds selected from the original countermodel. We first introduce a general construction together with the conditions required for it to yield a countermodel, and then instantiate this construction for each Åqvist logic, adapting it to the corresponding class of preferential models.

The resulting countermodel will contain a certain subset of worlds from the original countermodel M , some of them possibly copied multiple times, preserving the valuation from M but equipped with a new preference relation. We call transformations of this kind *rearrangements* of a model.

Definition 10. *We say that a model $\langle W', \succeq', V' \rangle$ rearranges the model $\langle W, \succeq, V \rangle$ when there exists a prototype function $prot: W' \rightarrow W$ such that $V'(w') = V(w)$.*

Our main goal for the rearranged model is to have each of its worlds satisfying the same subformulas of A as its prototype does. Evaluation of a formula in a world involves other worlds only in the cases of $\Box$ and $\bigcirc(\cdot|\cdot)$ modalities. Therefore, we only need to ensure that the rearranged model validates the same modalities among subformulas of A as the original model does, while the satisfaction (and non-satisfaction) of other subformulas will be preserved in the rearranged model automatically.

We will examine the cases of validated and non-validated modalities separately. Let us denote by $Box^+(A, M)$ (resp. $Ob^+(A, M)$) the set of subformulas of A of the form $\Box B$ (resp. $\bigcirc(D|C)$) that are validated by M , and by $Box^-(A, M)$ and $Ob^-(A, M)$ the sets of subformulas of A of the corresponding form that are not validated by M . To falsify $\Box B \in Box^-(A, M)$ and $\bigcirc(D|C) \in Ob^-(A, M)$ we need to take in M' some worlds that were falsifying these modalities in M . While the evaluation of $\Box B$ modalities relies only on the presence of the worlds satisfying B in the model, special care is needed to ensure that the evaluation of $\bigcirc(D|C)$ is the same. Namely, if a world w was made not C -maximal in M by some world $u \in \|C\|^M$ such that $u \succ w$ we need to preserve this violation of maximality in M' . Conversely, we need to ensure that we are not violating C -maximality in M' for the world falsifying $\bigcirc(D|C)$. This reasoning leads to the following four conditions sufficient to ensure that a rearranged model M' is a countermodel for A .

Theorem 9. *Suppose that $M' = \langle W', \succ', V' \rangle$ rearranges $M = \langle W, \succ, V \rangle$ with the prototype function $prot: W' \rightarrow W$, and $M \models A$. Then the following four conditions are sufficient for $M' \not\models A$.*

- (i) *There exists $v' \in W'$ such that $M, prot(v') \not\models A$.*
- (ii) *For any $\Box B \in Box^-(A, M)$ there exists $v' \in W'$ such that $M, prot(v') \not\models B$.*
- (iii) *For any $\bigcirc(D|C) \in Ob^-(A, M)$ there is $v' \in W'$ s.t. $prot(v') \in \max_{\succ}(\|C\|^M)$ and $prot(v') \notin \|D\|^M$ and for all $u' \succ' v'$ holds $prot(u') \succ prot(v')$.*
- (iv) *For any $w' \in W'$, for all $\bigcirc(D|C) \in Ob^+(A, M)$ if there exists $u \succ prot(w')$ such that $M, u \models C$ then there exists $s' \succ' w'$ such that $M, prot(s') \models C$.*

Proof. We will prove a generalized statement: for any $w' \in W'$ and $X \in SubF(A)$, $M', w' \models X$ iff $M, prot(w') \models X$. $M' \not\models A$ follows from this generalized statement by the condition (i). The proof is by induction on X (we use the abbreviation IH(s) to refer to the inductive hypothesis(es)).

- $X = x \in Atm$. $w \in V(x)$ is equivalent to $prot(w) \in V'(x)$ by the definition of the prototype function.
- X has a Boolean connective on top. The statement follows directly from IHs for immediate subformulas.
- $X = \Box B$ and $X \in Box^+(A, M)$. $M \models \Box B$, so for all $w' \in W'$, $M, prot(w') \models B$, so by IH $M', w' \models B$ for all $w' \in W'$, so $M' \models \Box B$.
- $X = \Box B$ and $X \in Box^-(A, M)$. By (ii) there is $v' \in W'$ such that $M, prot(v') \not\models B$, so by IH $M', v' \not\models B$, so $M' \not\models \Box B$.

- $X = \bigcirc(D|C)$ and $X \in Ob^+(A, M)$. Take any $w' \in \max_{\succ'}(\|C\|^{M'})$. Then $prot(w') \in \max_{\succ}(\|C\|^M)$: $prot(w') \in \|C\|^M$ by IH, and there can be no s s.t. $s \succ prot(w')$ and $s \in \|C\|^M$ (otherwise there is $u' \succ' w'$ s.t. $M', u' \models C$ by (iv) and IH). Since $M \models \bigcirc(D|C)$, $prot(w') \in \max_{\succ}(\|C\|^M)$ implies that $M, prot(w') \models D$, which implies $M', w' \models D$ by IH. Thus $M' \models \bigcirc(D|C)$.
- $X = \bigcirc(D|C)$ and $X \in Ob^-(A, M)$. For the corresponding world $v' \in W'$ from (iii) we have $v' \notin \|D\|^{M'}$ (by IH and the choice of v') and $v' \in \max_{\succ'}(\|C\|^{M'})$ (since $v' \in \|C\|^{M'}$ by IH and the choice of v' , and for all $u' \succ' v'$ we have $u' \notin \|C\|^{M'}$ by (iii) and IH), so $M' \not\models \bigcirc(D|C)$. $\square$

Ensuring conditions (i) and (ii) is simple: we need to take arbitrary worlds from $(W \setminus \|A\|^M)$ and from $(W \setminus \|B\|^M)$ for each $\Box B \in Box^-(A, M)$. For this, we will use a *representative function* $rep: (2^W \setminus \{\emptyset\}) \rightarrow W$ that chooses an element $rep(S) \in S$ from any given non-empty subset S of W (thus, we use the axiom of choice explicitly in our construction). We will also need representatives of $(\max_{\succ}(\|C\|^M) \setminus \|D\|^M)$ for every $\bigcirc(D|C) \in Ob^-(A, M)$ for condition (iii). Let us denote the set of all such falsifying worlds $Fal(A, M)$.

Definition 11. (*Falsifying worlds*) For a model $M = \langle W, \succeq, V \rangle$ such that $M \not\models A$, $Fal(A, M) = rep(W \setminus \|A\|^M) \cup Fal^{\Box}(A, M) \cup Fal^{\bigcirc}(A, M)$, where

$$\begin{aligned} Fal^{\Box}(A, M) &= \{rep(W \setminus \|B\|^M) \mid \Box B \in Box^-(A, M)\}, \\ Fal^{\bigcirc}(A, M) &= \{rep(\max_{\succ}(\|C\|^M) \setminus \|D\|^M) \mid \bigcirc(D|C) \in Ob^-(A, M)\}. \end{aligned}$$

The rest of the rearranged model will be chosen to ensure the satisfaction of conditions (iii) and (iv). We will represent our small model constructions as *composite models*, assembled from *blocks*. A *block* B is a finite selection of worlds from M with some new preference relation on them (in our cases it will be either an empty relation, a strict linear order, or a universal relation).

Definition 12. (*Block*) A block on M is a tuple $\langle U, \succeq_U \rangle$ where $U \subseteq W(M)$ and $\succeq_U$ is a binary relation on U . We will use $W(B)$ to refer to the set of worlds in B .

A composite construction consists of the number of blocks with an additional preference relation on them. Each composite construction generates a model rearranging M , in which the new preference relation is given by combining the relation between blocks and the relations inside blocks. To allow multiple occurrences of the same block in the construction we define the composite construction using labels and a labeling function.

Definition 13. (*Composite construction*) A composite construction on M is a tuple $\langle L, \succeq_L, \mathcal{B} \rangle$ where L is a set of labels, $\succeq_L$ is a binary relation on L , and $\mathcal{B}$ is a labeling function that maps every label from L into a block on M . Each composite construction $\mathfrak{C} = \langle L, \succeq_L, \mathcal{B} \rangle$ on $M = \langle W, \succeq, V \rangle$ generates a model $gen(\mathfrak{C}) = \langle W^{gen}, \succeq^{gen}, V^{gen} \rangle$, where

- $W^{gen} = \{(l, w) \mid l \in L, w \in W(\mathcal{B}(l))\}$;
- $(l_1, w_1) \succeq^{gen} (l_2, w_2)$ iff either $l_1 \succeq_L l_2$ or both $l_1 = l_2$ and $w_1 \succeq_U w_2$ for $\mathcal{B}(l_1) = \langle U, \succeq_U \rangle$;
- $V^{gen}((l, w)) = V(w)$.

We can now simplify the conditions of Th. 9 for models generated by composite constructions. For conditions (i)-(iii) it is enough to have every world v from $Fal(A, M)$ in some block in the construction, such that this block does not have any $\succ_U$ -preferable worlds inside and all blocks $\succ_L$ -prefferable to it have only worlds from $Bet_{\succ}(v) = \{w \in W : w \succ v\}$. Also, we can ensure condition (iv) separately for each block by either ensuring it inside this block or having another block $\succ_L$ -preferred to it that has all worlds required in condition (iv). We express these condition using the following notions of block compatibility.

Definition 14. (*Block compatibility properties*). For a given model M and a formula A we define the following properties of blocks on M :

- Block $\langle U, \succeq_U \rangle$ is flat when $w_1 \not\succ_U w_2$ for any $w_1, w_2 \in U$.
- Block B' is (iii)-suitable for a block B when $W(B') \subseteq Bet_{\succ}(w)$ for any $w \in W(B)$.
- Block $\langle U, \succeq_U \rangle$ is (iv)-safe when for every $w \in U$ and every $C \in Cond(A)$, if $C \in Sat\mathcal{F}_M(Bet_{\succ}(w))$ then there is $w' \succ_U w$ such that $M, w' \models C$.
- Block B' (iv)-covers B when $Sat\mathcal{F}_M(Bet_{\succ}(w)) \cap Cond(A) \subseteq Sat\mathcal{F}_M(W(B'))$ for any $w \in W(B)$, where $Sat\mathcal{F}_M(U) = \{X \in \mathcal{L}_{Aq} \mid \exists u \in U : M, u \models X\}$

Using these properties, we adapt Th. 9 for composite constructions.

Theorem 10. Let $M \not\models A$ and $\mathfrak{C} = \langle L, \succeq_L, \mathcal{B} \rangle$ be a composite construction on M . For $gen(\mathfrak{C}) \not\models A$ it is sufficient that:

- For every falsifying world $v \in Fal(A, M)$ there is a label $b_v \in L$ s.t. $v \in W(\mathcal{B}(b_v))$ and $\mathcal{B}(b_v)$ is flat and for every $b' \succ_L b_v$, $\mathcal{B}(b')$ is (iii)-suitable for $\mathcal{B}(b_v)$.
- For every $b \in L$ the block $\mathcal{B}(b)$ is either (iv)-safe or (iv)-covered by $\mathcal{B}(b')$ for some $b' \succ_L b$.

Proof. $gen(\mathfrak{C})$ rearranges M (with $prot((l, w)) = w$), so we can apply Th. 9 to it. Let $M = \langle W, \succeq, V \rangle$ and $gen(\mathfrak{C}) = \langle W', \succeq', V' \rangle$. We need to check conditions (i)-(iv) of Th. 9 for $\langle W', \succeq', V' \rangle$.

- Condition (i) is satisfied since there is $v \in Fal(A, M)$ such that $M, v \not\models A$, so by (a), $(l, v) \in W'$ for some label l .

- Proof of condition (ii) of Th. 9 is analogous.
- Condition (iii): For any $\bigcirc(D|C) \in Ob^-(A, M)$ there is $v \in Fal(A, M)$ s.t. both $v \in \max_{\succ}(\|C\|^M)$ and $v \notin \|D\|^M$, so there by (a) there exists a label $b_v \in L$ s.t. $v = prot(v')$ for some $v' \in W(\mathcal{B}(b_v))$. Since $\mathcal{B}(b_v)$ is flat by (a), $u' \succ' v'$ is true only if there is some $b' \in L$ s.t. $u' \in W(\mathcal{B}(b'))$ and $b' \succ_L b_v$. Therefore $prot(u') \succ prot(v')$ as required in condition (iii), since $\mathcal{B}(b')$ is (iii)-suitable for $\mathcal{B}(b_v)$ by (a).
- Condition (iv): For any $(l, w) \in W'$, and any $\bigcirc(D|C) \in Ob^+(A, M)$, suppose there exists $u \succ w$ such that $M, u \models C$. Let us consider the two cases given by (b) applied to block $\mathcal{B}(l)$. If $\mathcal{B}(l)$ is (iv)-safe then there is $w' \succ_U w$ (therefore $(l, w') \succ' (l, w)$) such $M, w' \models C$. If $\mathcal{B}(l)$ is (iv)-covered by $\mathcal{B}(l')$ for some $l' \in L$ such that $l' \succ_L l$, then $C \in Sat\mathcal{F}_M(W(\mathcal{B}(l')))$, since $C \in Sat\mathcal{F}_M(Bet_{\succ}(w))$, so $M, s \models C$ for some $(l', s) \in W'$ and $(l', s) \succ' (l, w)$. In either case, condition (iv) is satisfied. $\square$

Figure 4.2 provides an overview of the four resulting LEGO constructions. Their common idea is to retain a small collection of falsifying worlds and add only enough structure above them to preserve the truth of the relevant obligations. In the construction for **E**, each falsifying world v is represented by a singleton block fal_v and is followed by an orbit orb_v containing representatives of the relevant worlds better than v . Since **E** imposes no structural conditions on the preference relation, three copies of a common cloud are arranged in a cycle to provide the remaining witnesses required for the obligations. For **F**, limitedness allows these three cyclic cloud blocks to be replaced by a single finite chain ray , yielding an acyclic strict preference relation. For **F_{CM}**, transitivity allows us to go further: each falsifying world is paired with its own chain $c-ray_v$, so that the resulting model is essentially a union of incomparable finite chains. Finally, in **G**, totality requires these components to be coordinated globally. Falsifying worlds that are equivalent in the original total preorder are collected into clique blocks $group_{S_i}$; these are interleaved with the corresponding chains $g-ray_{S_i}$ and all blocks are linearly ordered. Thus, the increasingly structured constructions in Fig. 4.2 directly reflect the increasingly strong frame conditions of **E**, **F**, **F_{CM}**, and **G**.

4.4.1 LEGO construction for logic E

In the case of logic **E**, there are no model conditions we need to satisfy in the resulting countermodel, so we can use a preference relation that is non-transitive and contains cycles. In this simple case, all blocks of the countermodel construction will be antichains.

Definition 15. $Antichain(U) = \langle U, \succeq_a \rangle$ where $\succeq_a$ is the empty relation.

We start our composite construction for **E** with a dedicated one-world block $Antichain(\{v\})$ labeled fal_v for each world $v \in Fal(A, M)$. The simplest way to (iv)-cover such block with a (iii)-suitable block containing linearly many (w.r.t. $|A|$) worlds is to go through formulas from $Cond(A)$ satisfied in $Bet_{\succ}(v)$ and select one representative for each. Below,

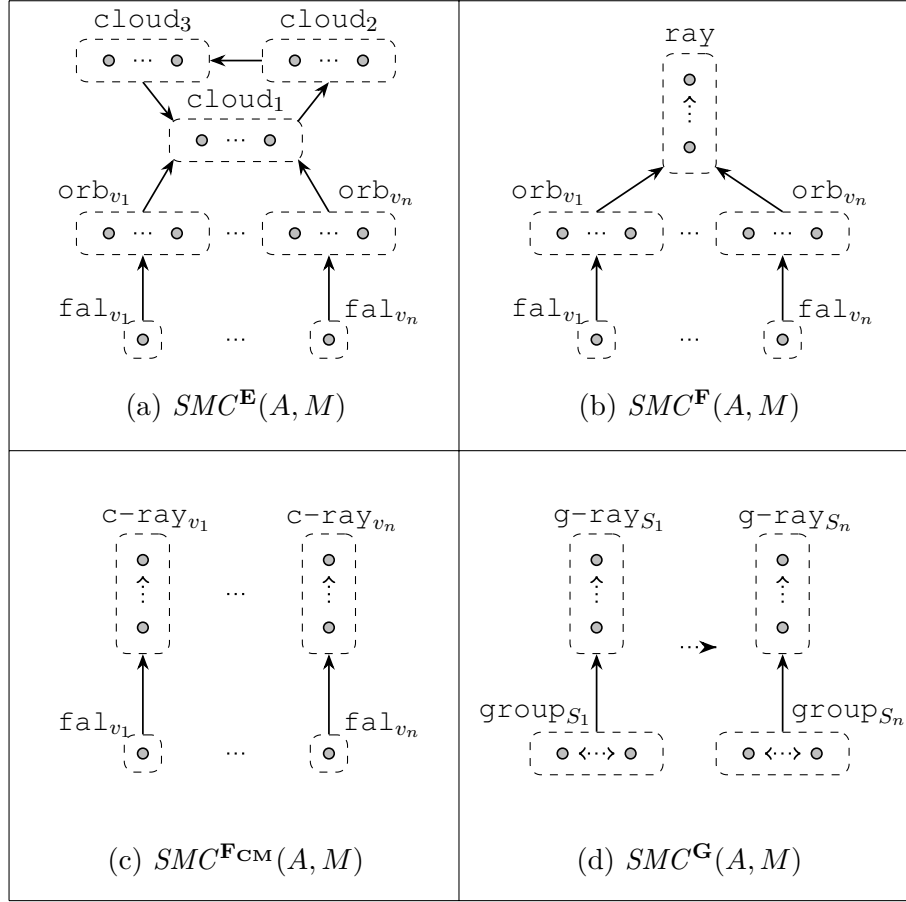


Figure 4.2: LEGO small model constructions for Åqvist's logics. Gray circles represent worlds, dashed rectangles represent blocks. Symbol $\dots$ inside a block indicates an antichain, $\hat{\cdot}$ indicates a chain, and $\langle \dots \rangle$ indicates a clique. Solid arrows represent the preference relation $\succeq_L$ between blocks: an arrow from a block l_1 to a block l_2 means $l_2 \succeq_L l_1$. The arrow $\dots \blacktriangleright$ between blocks in construction $SMC^G(A, M)$ means that there is a linear order on blocks. Note that the preference relation in constructions $SMC^E(A, M)$ and $SMC^F(A, M)$ is not transitive.

such a selection is defined more generally, for an arbitrary set of formulas $\mathcal{A}$ and an arbitrary set of worlds U to select from.

Definition 16. (*Selection*) For a set $U \subseteq W(M)$ and a set of formulas $\mathcal{A}$,

$$Sel_M(U, \mathcal{A}) = \{rep(\|C\|^M \cap U) \mid C \in \mathcal{A}, \|C\|^M \cap U \neq \emptyset\}.$$

We can show that such a selection (iv)-covers not only single-world blocks like fal_v , but any block B as long as U contains all worlds preferable to worlds in B .

Lemma 5. *If $Bet_{\succ}(w) \subseteq U$ for all $w \in W(B)$ in some block B then B is (iv) -covered by $Antichain(Sel_M(U, Cond(A)))$.*

Proof. If $C \in Cond(A)$ is satisfiable in $Bet_{\succ}(w)$ for some $w \in W(B)$ then there will be a representative satisfying C in $Sel_M(U, Cond(A))$. $\square$

A block to (iv) -cover $fail_v$, which we will call *orbit* and label orb_v , can be defined as

$$Orbit(M, A, v) = Antichain(Sel_M(Bet_{\succ}(v), Cond(A))).$$

To (iv) -cover orbits themselves, we can make another selection, this time from the whole $W(M)$, as this block does not need to be (iii) -suitable. So the block

$$Cloud(M, A) = Antichain(Sel_M(W(M), Cond(A))).$$

which we will label $cloud_1$, can be used to (iv) -cover all orbits. Finally, to (iv) -cover $cloud_1$ we can add two more copies of $Cloud(M, A)$ (labeled $cloud_2$ and $cloud_3$) and have a non-transitive loop on these three copies, which will (iv) -cover each other circularly. This leads to the following small model construction for **E**.

Definition 17. (*LEGO Construction for E*) Suppose M is an **E**-countermodel for A . $SMC^E(A, M) = \langle L, \succeq_L, \mathcal{B} \rangle$ where

$$\begin{aligned} L &= \{fail_v, orb_v \mid v \in Fal(A, M)\} \cup \{cloud_i \mid i \in \{1, 2, 3\}\}, \\ \mathcal{B}(fail_v) &= Antichain(\{v\}), \\ \mathcal{B}(orb_v) &= Orbit(M, A, v), \\ \mathcal{B}(cloud_i) &= Cloud(M, A), \end{aligned}$$

and the preference relation $\succeq_L$ on blocks is demonstrated on Fig. 4.2a.

Theorem 11. *If M is an **E**-countermodel for A then $gen(SMC^E(A, M))$ is also an **E**-countermodel for A and $|W(gen(SMC^E(A, M)))| = \mathcal{O}(|A|^2)$.³*

Proof. $SMC^E(A, M)$ is a countermodel for A by Th. 10, because $\mathcal{B}(fail_v)$ is flat, $\mathcal{B}(orb_v)$ is (iii) -suitable for $\mathcal{B}(fail_v)$, and all blocks are (iv) -covered by Lem. 5. Additionally, $|W(gen(SMC^E(A, M)))| = \mathcal{O}(|A|^2)$ since there are $(2 \cdot |Fal(A, M)| + 3)$ blocks $SMC^E(A, M)$ with at most $|Cond(A)|$ worlds in each block. $\square$

Example 4. Consider the formula

$$A = \bigcirc(\perp \mid a) \rightarrow \Box \neg a,$$

³As usual, the notation $f(A, M) = \mathcal{O}(g(A, M))$ for integer-valued functions f and g means that there exists a constant C such that $f(A, M) \leq C \cdot g(A, M)$ for all A and M .

an instance of $(D*)$ axiom, which is valid in $\mathbf{F}$ but not in $\mathbf{E}$. Let's consider the infinite countermodel M , containing an ascending chain of worlds

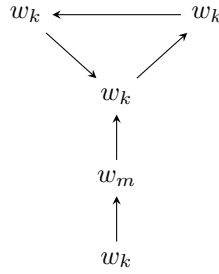
$$w_0 \prec w_1 \prec w_2 \prec \dots$$

in which every world satisfies a . Since there is no best a -world, $\bigcirc(\perp \mid a)$ holds vacuously, while $\Box\neg a$ is false, and hence $M \not\models A$.

Let $w_k = \text{rep}(W)$. Since $\text{Cond}(A) = \{a\}$, $\text{Box}^-(A, M) = \{\Box\neg a\}$ and $\text{Ob}^-(A, M) = \emptyset$, we have $\text{Fal}^\Box(A, M) = \text{Fal}(A, M) = \{w_k\}$ and $\text{Fal}^\bigcirc(A, M) = \emptyset$. The corresponding blocks in the construction are therefore

$$\begin{aligned} \mathcal{B}(\text{fal}_{w_k}) &= \text{Antichain}(\{w_k\}), \\ \mathcal{B}(\text{orb}_{w_k}) &= \text{Orbit}(M, A, w_k) = \text{Antichain}(\{w_m\}), \quad \text{where } w_m = \text{rep}(\{w_i\}_{i=k+1}^\infty) \\ \mathcal{B}(\text{cloud}_i) &= \text{Cloud}(M, A) = \text{Antichain}(\{w_k\}) \end{aligned}$$

Thus the infinite ascending chain is transformed into the following non-transitive LEGO-countermodel $\text{gen}(\text{SMC}^\mathbf{E}(A, M))$ (here arrow from w to w' denotes $w \preceq w'$):



In $\text{gen}(\text{SMC}^\mathbf{E}(A, M))$, every world satisfies a , while the three cloud copies form a cycle, so there are no best a -worlds. Hence $\bigcirc(\perp \mid a)$ holds vacuously, whereas $\Box\neg a$ is false. Therefore $\text{gen}(\text{SMC}^\mathbf{E}(A, M)) \not\models A$.

4.4.2 LEGO construction for logic $\mathbf{F}$

For $\mathbf{F}$, we will utilize the limitedness of the initial countermodel M to construct a small countermodel with an acyclic (strict) preference relation, which will automatically make the resulting countermodel limited too.

Lemma 6. *If W is finite and $\succ$ is acyclic then model $\langle W, \succeq, V \rangle$ is limited.*

Proof. If there is some $w_0 \in \|C\|^M$, consider (any) longest path $w_0 \prec w_1 \prec w_2 \prec \dots$ with worlds from $\|C\|^M$ starting from w_0 . Since W is finite and there can be no repetitions on the path (due to acyclicity of $\succ$), the path is finite and there is the last world w_m for which there is no $u \in \|C\|^M$ such that $u \succ w_m$, and so $w_m \in \max_\succ(\|C\|^M)$ by definition. $\square$

For acyclicity, we modify our construction $SMC^{\mathbf{B}}(A, M)$ by replacing a non-transitive cycle of blocks $\text{cloud}_1, \text{cloud}_2, \text{cloud}_3$ with a finite chain. To define the linear order in such chains formally, we use the notation of lists: $[a_1, \dots, a_n]$ denotes the list containing elements $\{a_i\}_{i=1}^n$ ordered by increasing indices, $[]$ denotes the empty list, and $a :: S$ denotes the list obtained by appending the element a to the beginning of the list S .

Definition 18. (*Decreasing chain*) For a list $S = [w_1, \dots, w_m]$ of worlds from $W(M)$, $\text{chain}(S) = \langle \{w_i\}_{i=1}^n, \succeq^{ch} \rangle$ where $w_i \succeq^{ch} w_j$ when $i \leq j$.

Our goal is to select a chain of polynomial size that is (iv) -safe and satisfies any condition $C \in \text{Cond}(A)$ that is satisfiable in M (which will allow us to use the chain to (iv) -cover any block). We construct such a chain through an iterative process, that selects maximal worlds for disjunctions of conditions. At the beginning of the process, we have set $\mathcal{A}_0 = \text{Cond}(C)$ as the set of conditions for which we need satisfying worlds. If at least one of conditions in $\mathcal{A}_0$ is satisfied by some world in M , then $\|\bigvee_{C \in \mathcal{A}_0} C\|^M \neq \emptyset$, then by limitedness there exists some $z_0 \in \max_{\succ}(\|\bigvee_{C \in \mathcal{A}_0} C\|^M)$. We can safely take z_0 as the first (i.e. most preferable) world in the chain, since there are no worlds $u \succ z_0$ in M satisfying conditions from $\mathcal{A}_0$. z_0 satisfies some conditions from $\mathcal{A}_0$ (since $M, z_0 \models \bigvee_{C \in \mathcal{A}_0} C$), therefore we can move on to the next step with a strictly smaller set $\mathcal{A}_1$ of conditions for which we still need satisfying worlds. We can safely repeat this process by taking worlds from $z_i \in \max_{\succ}(\|\bigvee_{C \in \mathcal{A}_i} C\|^M)$ at every iteration: z_i has no preferable C -worlds for all remaining conditions $C \in \mathcal{A}_i$, while for all already removed formulas there is a satisfying world somewhere earlier (i.e. preferable to z_i) in the chain, thus condition (iv) will be satisfied for this world. After a linear number of iterations, the chain will contain satisfying worlds for all formulas from $\mathcal{A}$ satisfiable in M .

Below is the formal definition of the described chain of maximal worlds. We give a generalized version that selects maximal worlds from any given subset of worlds U and any given set of formulas $\mathcal{A}_i$, the same way as we did for $\text{Sel}_M(U, \mathcal{A})$. We will need this generalized version for constructions for $\mathbf{F}_{\mathbf{CM}}$ and $\mathbf{G}$ too.

Definition 19. For any $U \subseteq W(M)$ and a finite set of formulas $\mathcal{A}_i$, the maximal sequence $\text{MaxSeq}_M(U, \mathcal{A}_i)$ is defined as follows:

$$\text{MaxSeq}_M(U, \mathcal{A}_i) = \begin{cases} [], & \text{if } \mathcal{A}_i = \emptyset \\ [], & \text{if } \mathcal{D}(U, \mathcal{A}_i) = \emptyset \\ d(U, \mathcal{A}_i) :: \text{MaxSeq}_M(U, \mathcal{A}_{i+1}), & \text{otherwise} \end{cases}$$

where

$$\begin{aligned} \mathcal{D}(U, \mathcal{A}) &= U \cap \max_{\succ}(\|\bigvee_{C \in \mathcal{A}} C\|^M), \\ d(U, \mathcal{A}) &= \text{rep}(\mathcal{D}(U, \mathcal{A})), \\ \mathcal{A}_{i+1} &= \{C \in \mathcal{A}_i \mid M, d(U, \mathcal{A}) \not\models C\}. \end{aligned}$$

Notice that for a finite $\mathcal{A}_0$ this sequence is well-defined (representative $d(U, \mathcal{A}_i)$ is always taken from a non-empty set and $|\mathcal{A}_i|$ decreases) and always has length at most $|\mathcal{A}_0|$. The

reasoning above that shows (iv)-safeness of the chain built from this sequence works in the general case with arbitrary U and does not even require the limitedness of M .

Lemma 7. *chain(MaxSeq_M(U, Cond(A))) is (iv)-safe for any $U \subseteq W(M)$.*

Proof. Suppose $z_k \in \text{MaxSeq}_M(U, \text{Cond}(A))$ and $\bigcirc(D|C) \in \text{Ob}^+(A, M)$. Then for some step k and set $\mathcal{A}_k$ of remaining conditions, $z_k \in \max_{\succ}(\|\bigvee_{C \in \mathcal{A}_k} C\|^M)$. If there is $u \succ z_k$ such that $M, u \models C$, then $C \notin \mathcal{A}_k$ due to maximality of z_k , which means that C was removed at some previous step, therefore there is z_j with $j < k$ such that $M, z_j \models C$. $\square$

For logic **F**, we select worlds in the chain from the whole $W(M)$: for a limited model M we define block

$$\text{Ray}(M, A) = \text{chain}(\text{MaxSeq}_M(\text{Cond}(A), W(M))),$$

which we will label ray . $\text{Ray}(M, A)$ contains satisfying worlds for all conditions from $\text{Cond}(A)$ satisfiable in M so it (iv)-covers any block on M .

Lemma 8. *For a limited M , $\text{Ray}(M, A)$ (iv)-covers any block.*

Proof. Let $|\text{MaxSeq}_M(\text{Cond}(A), W(M))| = m$. If a condition $C \in \text{Cond}(A)$ is satisfiable in M then C can not belong to the set $\mathcal{A}_m$ of the remaining conditions not covered by the maximal sequence, otherwise $\mathcal{A}_m \neq \emptyset$ and $\mathcal{D}(W(M), \mathcal{A}_m) \neq \emptyset$ due to limitedness of M . Therefore $M, z \models C$ for some world z in $\text{Ray}(M, A)$. $\square$

Replacing the non-transitive triangle in $\text{SMC}^E(A, M)$ with $\text{Ray}(M, A)$ yields the small-model construction $\text{SMC}^F(A, M)$ with an acyclic strict preference relation.

Definition 20. (*LEGO Construction for F*) Suppose M is an **F**-countermodel for A . $\text{SMC}^F(A, M) = \langle L, \succeq_L, \mathcal{B} \rangle$ where

$$\begin{aligned} L &= \{\text{fal}_v, \text{orb}_v \mid v \in \text{Fal}(A, M)\} \cup \{\text{ray}\}, \\ \mathcal{B}(\text{fal}_v) &= \text{Antichain}(\{v\}), \\ \mathcal{B}(\text{orb}_v) &= \text{Orbit}(M, A, \text{Bet}_{\succ}(v)), \\ \mathcal{B}(\text{ray}) &= \text{Ray}(M, A), \end{aligned}$$

and a preference relation $\succeq_L$ on blocks is demonstrated on Fig. 4.2b.

Theorem 12. *If M is an **F**-countermodel for A then $\text{gen}(\text{SMC}^F(A, M))$ is also an **F**-countermodel for A and $|W(\text{gen}(\text{SMC}^F(A, M)))| = \mathcal{O}(|A|^2)$.*

Proof. $\text{gen}(\text{SMC}^F(A, M))$ is a countermodel for A by Th. 10: $\mathcal{B}(\text{orb}_v)$ is (iii)-suitable for $\mathcal{B}(\text{fal}_v)$, $\mathcal{B}(\text{ray})$ is (iv)-safe by Lem. 7 and all other blocks are (iv)-covered by Lem. 8 and Lem. 5. Moreover, $\text{gen}(\text{SMC}^F(A, M))$ is an **F**-countermodel by Lem. 6. $|W(\text{gen}(\text{SMC}^F(A, M)))| = \mathcal{O}(|A|^2)$ since $\text{SMC}^F(A, M)$ contains $(2 \cdot |\text{Fal}(A, M)| + 1)$ blocks with at most $|\text{Cond}(A)|$ worlds in each block. $\square$

Example 5. Consider the formula

$$A = (\bigcirc(b \mid a) \wedge \bigcirc(c \mid a)) \rightarrow \bigcirc(c \mid a \wedge b),$$

an instance of the (CM) axiom, which is valid in $\mathbf{F}_{\mathbf{CM}}$ but not in $\mathbf{F}$. Let's consider the infinite $\mathbf{F}$ -countermodel M consisting of two incomparable chains

$$u_0 \prec u_1 \prec u_2 \prec \cdots \quad \text{and} \quad v_0 \prec v_1,$$

with

$$V(u_0) = \{a, b\}, \quad V(v_1) = \{a, b, c\}, \quad V(w) = \{a\} \text{ for } w \notin \{u_0, v_1\}.$$

The model is limited. Moreover, v_1 is the only best a -world, so both $\bigcirc(b \mid a)$ and $\bigcirc(c \mid a)$ hold. On the other hand, both u_0 and v_1 are best $a \wedge b$ -worlds, and $u_0 \not\models c$. Hence $\bigcirc(c \mid a \wedge b)$ is false and $M \not\models A$.

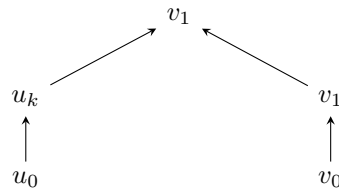
In this example we suppose $\text{rep}(W) = v_0$. Since $\text{Cond}(A) = \{a, a \wedge b\}$, $\text{Box}^-(A, M) = \emptyset$, and $\text{Ob}^-(A, M) = \{\bigcirc(c \mid a \wedge b)\}$, we have $\text{Fal}^\square(A, M) = \emptyset$, $\text{Fal}^\bigcirc(A, M) = \{u_0\}$, $\text{Fal}(A, M) = \{v_0, u_0\}$.

The corresponding blocks in the construction are therefore

$$\begin{aligned} \mathcal{B}(\text{fal}_{v_0}) &= \text{Antichain}(\{v_0\}), \\ \mathcal{B}(\text{fal}_{u_0}) &= \text{Antichain}(\{u_0\}), \\ \mathcal{B}(\text{orb}_{v_0}) &= \text{Orbit}(M, A, v_0) = \text{Antichain}(\{v_1\}) \\ \mathcal{B}(\text{orb}_{u_0}) &= \text{Orbit}(M, A, u_0) = \text{Antichain}(\{u_k\}) \quad \text{where } u_k = \text{rep}(\{u_i\}_{i=1}^\infty) \\ \mathcal{B}(\text{ray}) &= \text{Ray}(M, A) = \text{chain}([v_1]) \end{aligned}$$

Indeed, v_1 is the unique best world satisfying $a \vee (a \wedge b)$ and already satisfies both conditions from $\text{Cond}(A)$, so the maximal sequence defining $\text{Ray}(M, A)$ terminates after selecting v_1 .

Thus the infinite countermodel M is transformed into the following non-transitive LEGO-countermodel $\text{gen}(\text{SMC}^\mathbf{F}(A, M))$:



In $\text{gen}(\text{SMC}^\mathbf{F}(A, M))$, the unique best a -world is the top-most v_1 , so both $\bigcirc(b \mid a)$ and $\bigcirc(c \mid a)$ hold. The best $a \wedge b$ -worlds are u_0 and the top-most v_1 ; since $u_0 \not\models c$, $\bigcirc(c \mid a \wedge b)$ is false. Hence $\text{gen}(\text{SMC}^\mathbf{F}(A, M)) \not\models A$.

4.4.3 LEGO construction for logic $\mathbf{F}_{\mathbf{CM}}$

For $\mathbf{F}_{\mathbf{CM}}$ we need to ensure the transitivity of the preference relation in the transformed countermodel. It is enough to obtain an $\mathbf{F}_{\mathbf{CM}}$ -countermodel since for finite models transitivity implies smoothness.

Lemma 9. *$M = \langle W, \succeq, V \rangle$ is smooth if W is finite and $\succeq$ is transitive.*

Proof. First, notice that transitivity of $\succeq$ implies transitivity of $\succ$. Indeed, if $w_1 \succ w_2$ and $w_2 \succ w_3$ then $w_1 \succeq w_3$ by transitivity of $\succeq$ and $w_3 \not\succ w_1$ since otherwise there would be a transitive triangle on these three worlds and none of them could be strictly preferable to another. Now, for an arbitrary $w_0 \in ||C||^M$, consider (any) longest path $w_0 \prec w_1 \prec w_2 \prec \dots$ with worlds from $||C||^M$ starting from w_0 . Since W is finite and there can be no repetitions on the path due to transitivity of $\succ$, the path is finite and there is the last world w_m for which there is no $u \in ||C||^M$ such that $u \succ w_m$, so $w_m \in \max_{\succ}(||C||^M)$ and either $w_m = w_0$ or $w_m \succ w_0$. Therefore, M is smooth. $\square$

In $SMC^{\mathbf{F}}(A, M)$ non-transitivity was essential: we can not assert $\text{ray} \succ_L \text{fal}_v$ since we selected worlds in the maximal chain from the whole initial model, so $\mathcal{B}(\text{ray})$ can be not (iii)-suitable for $\mathcal{B}(\text{fal}_v)$. However, smoothness allows us to select a maximal chain only among worlds in $\text{Bet}_{\succ}(v)$. Specifically, for every falsifying world v we introduce individual *chain-orbit*

$$\text{RayOrb}(M, A, v) = \text{chain}(\text{MaxSeq}_M(\text{Bet}_{\succ}(v), \text{Cond}(A))).$$

We already know that this block is (iv)-safe by Lem. 7, and we can show that for an $\mathbf{F}_{\mathbf{CM}}$ -model M it covers block $\mathcal{B}(\text{fal}_v)$.

Lemma 10. *If M is transitive and smooth, $\text{RayOrb}(M, A, v)$ (iv)-covers $\text{Antichain}(\{v\})$.*

Proof. Suppose that (1) there is some $u \succ v$ in M such that $M, u \models C$ for some $C \in \text{Cond}(A)$, we need to show that there is a world z_k in the chain such that $M, z_k \models C$. Similarly to Lem. 8, we show it by proving that in this case C is removed from the set of conditions $\mathcal{A}_i$ at some point. And to show this, it is enough to prove that for smooth models (*) $C \in \mathcal{A}_i$ implies $\mathcal{D}(\text{Bet}_{\succ}(v), \mathcal{A}_i) \neq \emptyset$ from the definition of MaxSeq (then the sequence of maximal worlds cannot end while C belongs to $\mathcal{A}_i$).

Let us prove (*). Suppose that $C \in \mathcal{A}_i$. From this and (1) follows $u \in ||\bigvee_{C \in \mathcal{A}_i} C||^M$. Due to smoothness of M it implies that (2) there is $u' \in \max_{\succ}(||\bigvee_{C \in \mathcal{A}_i} C||^M)$ such that either $u' = u$ or $u' \succ u$. In either case $u' \succ v$ (since $u \succ v$ by (1) and transitivity of $\succeq$ implies transitivity of $\succ$). So, we have (3) $u' \in \text{Bet}_{\succ}(v)$. (2) and (3) together imply $u' \in \mathcal{D}(\text{Bet}_{\succ}(v), \mathcal{A}_i)$, concluding the proof of (*). $\square$

So we can obtain a small model construction for $\mathbf{F}_{\mathbf{CM}}$ by replacing each orbit orb_v in $SMC^{\mathbf{F}}(A, M)$ with an individual maximal chain $\text{RayOrb}(M, A, v)$ (which we will label c-ray_v). The common chain ray from $SMC^{\mathbf{F}}(A, M)$ is not needed anymore.

Definition 21. (*LEGO Construction for $\mathbf{F}_{\mathbf{CM}}$*) Suppose M is an $\mathbf{F}$ -countermodel for A . $SMC^{\mathbf{F}_{\mathbf{CM}}}(A, M) = \langle L, \succeq_L, \mathcal{B} \rangle$ where

$$\begin{aligned} L &= \{fal_v, c-ray_v \mid v \in Fal(A, M)\}, \\ \mathcal{B}(fal_v) &= Antichain(\{v\}), \\ \mathcal{B}(c-ray_v) &= RayOrb(M, A, v), \end{aligned}$$

and a preference relation $\succeq_L$ on blocks is demonstrated on Fig. 4.2c.

Theorem 13. If M is an $\mathbf{F}_{\mathbf{CM}}$ -countermodel for A then $gen(SMC^{\mathbf{F}_{\mathbf{CM}}}(A, M))$ is also an $\mathbf{F}_{\mathbf{CM}}$ -countermodel for A and $|W(gen(SMC^{\mathbf{F}_{\mathbf{CM}}}(A, M)))| = \mathcal{O}(|A|^2)$.

Proof. $SMC^{\mathbf{F}_{\mathbf{CM}}}(A, M)$ is a countermodel for A by Th. 10: $\mathcal{B}(c-ray_v)$ is (iii)-suitable for $\mathcal{B}(fal_v)$, each $\mathcal{B}(c-ray_v)$ is (iv)-safe by Lem. 7 and each $\mathcal{B}(fal_v)$ is (iv)-covered by $\mathcal{B}(c-ray_v)$ by Lem. 10. $SMC^{\mathbf{F}_{\mathbf{CM}}}(A, M)$ is an $\mathbf{F}_{\mathbf{CM}}$ -countermodel by Lem. 9. Finally, $|W(gen(SMC^{\mathbf{F}_{\mathbf{CM}}}(A, M)))| = \mathcal{O}(|A|^2)$ since $SMC^{\mathbf{F}_{\mathbf{CM}}}(A, M)$ contains $(2 \cdot |Fal(A, M)|)$ blocks with at most $|Cond(A)|$ worlds each. $\square$

Example 6. Consider the formula

$$A = (\neg \bigcirc (\neg b \mid a) \wedge \bigcirc (c \mid a)) \rightarrow \bigcirc (c \mid a \wedge b),$$

an instance of the (RM) axiom, which is valid in $\mathbf{G}$ but not in $\mathbf{F}_{\mathbf{CM}}$. Let's consider the infinite countermodel M consisting of an infinite chain

$$u_0 \prec u_1 \prec u_2 \prec \cdots \prec u_\omega \prec u_{\omega+1},$$

and an isolated world v , with

$$V(u_0) = \{a, b\}, \quad V(v) = \{a, b, c\}, \quad V(u_{\omega+1}) = \{a, c\},$$

and

$$V(u_i) = \{a\} \quad \text{for } 0 < i \leq \omega.$$

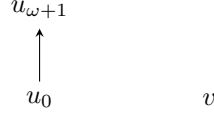
The model is transitive and smooth. Its best a -worlds are $u_{\omega+1}$ and v . Hence $\bigcirc(c \mid a)$ holds, while $\bigcirc(\neg b \mid a)$ does not, since $v \models b$. On the other hand, the best $a \wedge b$ -worlds are u_0 and v , and $u_0 \not\models c$. Thus $\bigcirc(c \mid a \wedge b)$ is false and $M \not\models A$.

Suppose for this example that $rep(W) = v$. Since $Cond(A) = \{a, a \wedge b\}$, $Box^-(A, M) = \emptyset$, $Ob^-(A, M) = \{\bigcirc(\neg b \mid a), \bigcirc(c \mid a \wedge b)\}$, we have $Fal^\square(A, M) = \emptyset$, and $Fal^\bigcirc(A, M) = Fal(A, M) = \{v, u_0\}$. The corresponding blocks in the construction are therefore

$$\begin{aligned} \mathcal{B}(fal_v) &= Antichain(\{v\}), \\ \mathcal{B}(fal_{u_0}) &= Antichain(\{u_0\}), \\ \mathcal{B}(c-ray_v) &= RayOrb(M, A, v) = chain(\square), \\ \mathcal{B}(c-ray_{u_0}) &= RayOrb(M, A, u_0) = chain([u_{\omega+1}]). \end{aligned}$$

Indeed, v is isolated, so its chain-orbit is empty. Above u_0 , only the condition a is satisfiable, and $u_{\omega+1}$ is its best world; hence the maximal sequence defining the second chain-orbit consists only of $u_{\omega+1}$.

Thus the infinite countermodel is transformed into the following LEGO-countermodel $\text{gen}(\text{SMC}^{\mathbf{FCM}}(A, M))$ (here an arrow from w to w' denotes $w \preceq w'$):



In $\text{gen}(\text{SMC}^{\mathbf{FCM}}(A, M))$, the best a -worlds are $u_{\omega+1}$ and v , so $\neg \bigcirc (\neg b \mid a)$ and $\bigcirc (c \mid a)$ hold. The best $a \wedge b$ -worlds are u_0 and v ; since $u_0 \not\models c$, $\bigcirc (c \mid a \wedge b)$ is false. Hence $\text{gen}(\text{SMC}^{\mathbf{FCM}}(A, M)) \not\models A$.

Remark 2. The form of the countermodel that we obtain — a union of incomparable finite chains — is the same as a Friedman-Halpern countermodel for logic **PCLTA** (i.e. **FCM**) [FH94]. However, we have achieved it by using different methods: they use a finite-model property of **PCL** extensions (shown in [Bur81]) and extend the preference relation to a linear order, then construct chains by selecting the greatest world w.r.t. extended order independently for each conditional in $\text{Ob}^+(A, M)$, while we do it using an iterative procedure. The possibility of their selection fully relies on finiteness and transitivity, which due to Lem. 9 is only possible in smooth models, so it cannot be applied to the weaker logics **E** and **F**. Furthermore, the conditional entailment⁴ in **PCLTA** was studied extensively in the area of non-monotonic reasoning, as it coincides with the KLM logic **P** [KLM90] of preferential reasoning (see Chapter 3). A small model construction for **P** has been introduced in [LM92] and consists of a single chain of polynomial size (using essentially the same method as Friedman and Halpern). Notice that both Friedman-Halpern and our constructions turn into a single chain when restricted to the conditional entailment.

4.4.4 LEGO construction for logic **G**

For logic **G**, we also need to ensure the totality of the preference relation in the transformed model by leveraging the fact that the falsifying worlds in $\text{Fal}(A, M)$ are ordered in the initial model by $\succeq$ which is a total preorder.

Let us consider first a simple case where $\succeq$ in the given **G**-countermodel is asymmetric (and therefore a strict linear order). Then there exists an ordering $v_1 \prec \dots \prec v_n$ of worlds from $\text{Fal}(A, M)$. Then we can linearly order blocks of $\text{SMC}^{\mathbf{FCM}}(A, M)$ with the following order: $\text{fal}_{v_1} \prec_L \text{c-ray}_{v_1} \prec_L \dots \prec_L \text{fal}_{v_n} \prec_L \text{c-ray}_{v_n}$. The (iii)-suitability will still

⁴Conditional entailment is a formula of the form $\bigcirc(D_1|C_1) \wedge \dots \wedge \bigcirc(D_n|C_n) \rightarrow \bigcirc(D_0|C_0)$, where C_i and D_i do not contain modalities.

be satisfied with such ordering, because for every $j \geq i$ we have $W(\text{c-ray}_{v_j}) \subseteq \text{Bet}_{\succ}(v_j)$ and $\text{Bet}_{\succ}(v_j) \subseteq \text{Bet}_{\succ}(v_i)$ due to transitivity of $\succeq$.

In general, the preference relation $\succeq$ in $\mathbf{G}$ -models is not necessarily asymmetric, but we can generalize the same idea by grouping together $\succeq$ -equivalent worlds as in the following definition.

Definition 22. (*Stratification*) For $M = \langle W, \succeq, V \rangle$ and $U \subseteq W$, a sequence $[S_1, \dots, S_n]$ of non-empty subsets of W is called a stratification of U when U is the disjoint union of subsets $\{S_i\}_{i=1}^n$ and for every $u_i \in S_i, u_j \in S_j$, $s_i \succeq s_j$ iff $i \geq j$.

When $\succeq$ is a total preorder, the unique stratification of any finite subset of worlds U is given by its factorization w.r.t. $\succeq$ -equivalence.

Lemma 11. *If $\succeq$ is transitive and total, there exists a unique stratification of any finite subset U .*

Proof. Consider an equivalence relation $\approx$ on U where $u_1 \approx u_2$ means that both $u_1 \succeq u_2$ and $u_2 \succeq u_1$. Consider further a relation $\succeq_S$ on the set of equivalence classes of U w.r.t. $\approx$ where $S_i \succeq_S S_j$ when there exist $u_i \in S_i$ and $u_j \in S_j$ such that $u_i \succeq u_j$. Notice that for a transitive and total $\succeq$ the relation $\succeq_S$ is a linear order: it is antisymmetric due to definitions of $\approx$ and $\succeq_S$, and it is transitive and total (and hence reflexive) due to the transitivity and totality of $\succeq$. This linear ordering gives a stratification by definition. Notice also that it is the only stratification: every element of a stratification should be an equivalence class w.r.t. $\approx$ and their order in the list should be aligned with $\succeq_S$ (i.e. $i \geq j$ implies $S_i \succeq_S S_j$) by definition. $\square$

This shows that the falsifying worlds $\text{Fal}(A, M)$ in M form a linearly ordered sequence of cliques (i.e. equivalence classes under $\succeq$). These cliques can then serve as the blocks in our composite model construction for $\mathbf{G}$, instead of the singleton falsifying worlds used previously.

Definition 23. (*Clique*) For $U \subseteq W(M)$, $\text{clique}(U) = \langle U, \succeq^{cl} \rangle$ where $u_1 \succeq^{cl} u_2$ for all $u_1, u_2 \in U$.

Notice that for $u_1, u_2 \in S_i$ both $u_1 \succeq u_2$ and $u_2 \succeq u_1$ so $\text{Bet}_{\succ}(u_1) = \text{Bet}_{\succ}(u_2)$ (due to transitivity). Therefore, to (iv)-cover block $\text{clique}(S_i)$ we can take orbit-chain $\text{RayOrb}(M, A, \text{rep}(S_i))$ with an arbitrary representative of S_i .

Lemma 12. *For a transitive and smooth $M = \langle W, \succeq, V \rangle$ and $U \subseteq W(M)$, if $u \succeq u'$ for all $u, u' \in U$ then $\text{RayOrb}(M, A, \text{rep}(U))$ (iv)-covers $\text{clique}(U)$.*

Proof. $\text{Bet}_{\succ}(u) = \text{Bet}_{\succ}(\text{rep}(U))$ for any $u \in U$ (since $\succeq$ is transitive), so any formula from $\text{Cond}(A)$ satisfied in some world from $\text{Bet}_{\succ}(u)$ is also satisfied by some world in $\text{RayOrb}(M, A, \text{rep}(U))$ by Lem. 10. $\square$

Thus we can take as the construction $SMC^{\mathbf{G}}(A, M)$ a linearly-ordered sequence of blocks in which cliques $\text{clique}(S_i)$, labeled group_{S_i} , are interleaved with chain-orbits $\text{RayOrb}(M, A, \text{rep}(S_i))$, labeled g-ray_{S_i} .

Definition 24. (*LEGO Construction for $\mathbf{G}$*) Suppose M is an $\mathbf{G}$ -countermodel for a formula A , and let $[S_1, \dots, S_n]$ be the unique stratification of $\text{Fal}(A, M)$ w.r.t. $\succeq$. Then $SMC^{\mathbf{G}}(A, M) = \langle L, \succeq_L, \mathcal{B} \rangle$ where

$$\begin{aligned} L &= \bigcup_{i=1}^n \{\text{group}_{S_i}, \text{g-ray}_{S_i}\}, \\ \mathcal{B}(\text{group}_{S_i}) &= \text{clique}(V_i), \\ \mathcal{B}(\text{g-ray}_{S_i}) &= \text{chain}(\text{RayOrb}(M, A, \text{rep}(S_i))), \end{aligned}$$

and the blocks are ordered linearly by $\succeq_L$ as follows:

$$\text{group}_{S_1} \prec_L \text{g-ray}_{S_1} \prec_L \dots \prec_L \text{group}_{S_n} \prec_L \text{g-ray}_{S_n}.$$

Theorem 14. If M is an $\mathbf{G}$ -countermodel for A then $\text{gen}(SMC^{\mathbf{G}}(A, M))$ is also an $\mathbf{G}$ -countermodel for A and $|W(\text{gen}(SMC^{\mathbf{G}}(A, M)))| = \mathcal{O}(|A|^2)$.

Proof. $SMC^{\mathbf{G}}(A, M)$ is a countermodel for A by Th. 10: $\mathcal{B}(\text{group}_{S_i})$ is flat and all blocks $\succ_L$ -preferable to it are (iii)-suitable for it (since $\succeq$ in M is transitive), each $\mathcal{B}(\text{g-ray}_{S_i})$ is (iv)-safe by Lem. 7, and each $\mathcal{B}(\text{group}_{S_i})$ is (iv)-covered by $\mathcal{B}(\text{g-ray}_{S_i})$ due to Lem. 12. $SMC^{\mathbf{G}}(A, M)$ is a $\mathbf{G}$ -countermodel since its preference relation is transitive and total, and therefore also smooth by Lem. 9. $|W(\text{gen}(SMC^{\mathbf{G}}(A, M)))| = \mathcal{O}(|A|^2)$ since there are at most $(2 \cdot |\text{Fal}(A, M)|)$ blocks in $SMC^{\mathbf{G}}(A, M)$ with at most $|\text{Cond}(A)|$ worlds in each block. $\square$

Example 7. Consider the formula

$$A = \bigcirc(b \mid a) \vee \bigcirc(\neg b \mid a),$$

an instance of conditional excluded middle, which is not valid in $\mathbf{G}$. Let's consider the infinite countermodel M consisting of an infinite descending sequence of preference levels

$$\{u_0, v_0\} \succ \{u_1, v_1\} \succ \{u_2, v_2\} \succ \dots,$$

where on each level we have $u_i \preceq v_i$ and $u_i \succeq v_i$, and

$$V(u_i) = \{a\}, \quad V(v_i) = \{a, b\} \quad \text{for every } i \geq 0.$$

The preference relation is a smooth total preorder. The best a -worlds are u_0 and v_0 , which disagree on b . Hence neither $\bigcirc(b \mid a)$ nor $\bigcirc(\neg b \mid a)$ holds, and therefore $M \not\models A$.

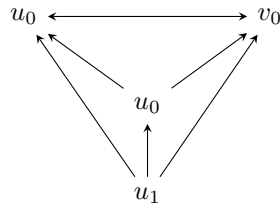
Suppose for this example $\text{rep}(W) = u_1$ and $\text{rep}(\{u_0, v_0\}) = u_0$. Since $\text{Cond}(A) = \{a\}$, $\text{Box}^-(A, M) = \emptyset$, $\text{Ob}^-(A, M) = \{\bigcirc(b \mid a), \bigcirc(\neg b \mid a)\}$, we have $\text{Fal}^\square(A, M) = \emptyset$, $\text{Fal}^\circ(A, M) = \{u_0, v_0\}$, $\text{Fal}(A, M) = \{u_1, u_0, v_0\}$. The stratification of $\text{Fal}(A, M)$

consists of $S_1 = \{u_1\}$, and $S_2 = \{u_0, v_0\}$. The corresponding blocks in the construction are therefore

$$\begin{aligned}\mathcal{B}(\text{group}_{S_1}) &= \text{clique}(\{u_1\}), \\ \mathcal{B}(\text{group}_{S_2}) &= \text{clique}(\{u_0, v_0\}), \\ \mathcal{B}(g\text{-ray}_{S_1}) &= \text{RayOrb}(M, A, u_1) = \text{chain}([u_0]), \\ \mathcal{B}(g\text{-ray}_{S_2}) &= \text{RayOrb}(M, A, u_0) = \text{chain}([]).\end{aligned}$$

Indeed, the worlds strictly preferable to u_1 are u_0 and v_0 , and both satisfy the only condition a , so the first chain-orbit contains their chosen representative u_0 . There are no worlds strictly preferable to u_0 or v_0 , hence the second chain-orbit is empty.

Thus the infinite countermodel is transformed into the following LEGO-countermodel $\text{gen}(\text{SMC}^{\mathbf{G}}(A, M))$:



In $\text{gen}(\text{SMC}^{\mathbf{G}}(A, M))$, the best a -worlds are u_0 and v_0 in the top clique. Since they disagree on b , both $\bigcirc(b \mid a)$ and $\bigcirc(\neg b \mid a)$ are false. Hence $\text{gen}(\text{SMC}^{\mathbf{G}}(A, M)) \not\models A$.

Remark 3. Friedman and Halpern also provide a counter-model for logic **VTa** (i.e. $\mathbf{G}$) [FH94]. They use an ad hoc approach, different from the one they use for the other extensions of **PCL**. For **VTa**, they simply take one world for each conditional from the original model without changing the preference relation, resulting in a model of linear size. Although Th. 9 can be also used to establish the adequacy of their construction, we provided a different construction (with a new explicitly defined preference relation and with potentially a quadratic number of worlds) for uniformity with the constructions for the other three Åqvist's logics.

4.5 Alternative frame conditions for Åqvist's logics

New semantical characterizations for theoremhood can be extracted from the specific form of our small model constructions. Namely, any model property satisfied by $\text{SMC}^L(A, M)$ that is stronger than some existing characterization for L (e.g. from Fig. 4.1) can be used as an alternative characterization.

Formally, we call a class $\mathfrak{M}$ of preference models a semantical characterization for theoremhood in logic L when any A is a theorem of L iff A is valid in all models from $\mathfrak{M}$.

Lemma 13. *Let $\mathfrak{M}$ be a class of models characterizing theoremhood in L . If $\mathfrak{M}' \subseteq \mathfrak{M}$ and $\text{SMC}^L(A, M) \in \mathfrak{M}'$ for every A and M , then $\mathfrak{M}'$ characterizes theoremhood in L .*

Logic	Cardinality of W	Properties of $\succeq$		
		acyclic	transitive	total
E	finite			
F	finite	✓		
F_{CM}	finite		✓	
G	finite		✓	✓

Table 4.1: Finite-model characterizations of theoremhood in Åqvist's logics (with maximality as the notion of bestness).

Proof. If A is a theorem of L , it is valid in all models in $\mathfrak{M}$, so it is also valid in all models of $\mathfrak{M}'$. If A is not a theorem of L , it is not valid in $SMC^L(A, M)$, which belongs to $\mathfrak{M}'$. $\square$

We can use this method to characterize theoremhood in Åqvist logics with *frame properties*, i.e. properties of the preference relation. Notice that the limit conditions (limitedness and smoothness) used for the characterization of **F**, **F_{CM}**, and **G** are not frame properties: they impose conditions only on truth sets of the model. This choice plays a vital role in establishing correspondence between semantics and known axiomatizations of Åqvist's logics, but it makes it hard to work with these models since you need to distinguish which subsets of worlds can be a truth set. However, we can notice that our small model constructions satisfy the corresponding limit conditions for all subsets of worlds, therefore limitedness/smoothness on the level of frames can be used to characterize theoremhood too.

More importantly, our constructions satisfy some stronger frame properties. We already used these properties to prove that $SMC^L(A, M)$ generates an L -countermodel. If we consider only finite models limit conditions can be replaced with natural conditions on preference relations: limitedness can be replaced with acyclicity, and smoothness can be dropped in presence of transitivity.

Theorem 15. *Formula A is a theorem of Åqvist logic L iff $M \models A$ for all finite models M satisfying the frame properties for logic L from Tab. 4.1.*

Proof. Using Lem. 13. Finite models with transitive preference relation are smooth by Lem. 9. Finite models with acyclic preference relation are limited by Lem. 6 (acyclicity of $\succeq$ implies acyclicity of $\succ$), and although the preference relation in $SMC^F(A, M)$ has cycles in the form of reflexive loops inside the chain, these loops can be removed without affecting the satisfaction in the model, so any non-theorem has a finite countermodel with acyclic preference relation. $\square$

In addition, our models for **E**, **F**, **F_{CM}** satisfy antisymmetry so this property can be added to finite-model characterization for these logics from Fig. 4.1, but not to the

characterization of $\mathbf{G}^5$. Either reflexivity or irreflexivity can also be added since it is trivial to force them in any model without changing the satisfaction relation. Thus, $\mathbf{F}_{\mathbf{CM}}$ is characterized by finite models where $\succeq$ is a partial order (strict or non-strict). At the same time, finite models where $\succeq$ is a linear order give some logic that is stronger than $\mathbf{G}$. Even more specialized properties can be extracted from our construction via Lem. 13, e.g. $\mathbf{F}_{\mathbf{CM}}$ can be characterized by models that are unions of non-comparable finite chains.

Remark 4. *Note that here we are only concerned with semantical characterizations of theoremhood. Our results can not be extended to characterizing entailments $\Gamma \vdash A$ if Γ is infinite (to provide a strongly complete characterization of the logics). This is a natural limitation for finite-model characterizations since entailments from infinite sets of premises can not be characterized using only finite models for $\mathbf{S5}$ already and therefore for Åqvist's logics that extend $\mathbf{S5}$ too. As an example, consider the infinite set of propositional variables $\{x_i\}_{i=1}^\infty$ and an infinite sequence of formulas $\{\varepsilon_n\}_{n=1}^\infty$ defined as*

$$\varepsilon_n = x_n \wedge \bigwedge_{i=1}^{n-1} \neg x_i$$

Clearly, two formulas ε_n and ε_m for $n \neq m$ can not be both satisfied in one world. Therefore, the formulas from the set $\Gamma^\diamond = \{\diamond \varepsilon_n\}_{n=1}^\infty$ can not all be simultaneously valid in any finite model. So entailment $\Gamma^\diamond \models \perp$ holds in all finite relational models, but does not hold in $\mathbf{S5}$, since an infinite relational model satisfying all formulas from $\Gamma^\diamond$ simultaneously can be easily constructed.

4.6 Complexity: Åqvist's logics are co-NP-complete

Our small model constructions show that for any non-valid formula there exists a countermodel with at most $N(\varphi)$ worlds, where $N(\varphi)$ is a certain upper bound polynomial w.r.t. $|\varphi|$. Plus, the stronger frame properties from Tab. 4.1 can be easily checked in polynomial time w.r.t. the model size. This immediately implies co-NP-completeness of theoremhood.

Theorem 16. *Theoremhood is co-NP-complete for every Åqvist's logic.*

Proof. Non-theoremhood can be checked non-deterministically in polynomial time by guessing a countermodel M of size at most $N(\varphi)$ (i.e. guessing preference relation and valuation for all variables occurring in φ) and then checking $M \not\models \varphi$ and the required properties from Fig. 4.1. co-NP-hardness follows from co-NP-completeness of theoremhood in classical logic (since a propositional formula is a classical tautology iff it is a theorem of an Åqvist logic). $\square$

⁵E.g., the principle of conditional excluded middle $\bigcirc(D|C) \vee \bigcirc(\neg D|C)$ is valid in all models with total and antisymmetric preference relation (since they have at most one C -best world), but it can be easily falsified in $\mathbf{G}$ by a model with two C -worlds preferable to each other (and thus both belonging to $\max_{\succ}(|C|^M)$ allowed to be C -best simultaneously).

4.7 Prover: Deo-SMT

The semantic characterization we introduced were employed for automated reasoning in Åqvist's logics in the prover Deo-SMT⁶ [Köl25]. The prover exploits the finite-model characterizations of the four Åqvist systems to obtain a modular SAT encoding. Given a formula, it constructs an SAT representation of a countermodel whose size is bounded by the corresponding small-model theorem: for **E**, **F**, and **F_{CM}** it uses the quadratic bounds obtained from our LEGO constructions, while for **G** it relies on the linear bound provided by the Friedman-Halpern construction. The encoding for the weakest logic **E** captures arbitrary finite preference models of the required size. Encodings for the stronger systems are then obtained incrementally by adding constraints corresponding to their characteristic frame conditions from Tab. 4.1: acyclicity for **F**, transitivity for **F_{CM}**, and totality and transitivity for **G**. In particular, acyclicity is represented by introducing an auxiliary relation encoding reachability through a non-empty sequence of preference steps and requiring this relation to be transitive, irreflexive, and to contain the original preference relation. The resulting architecture mirrors the modular semantic hierarchy of Åqvist's logics and allows all four systems to be handled uniformly within a single SMT-based framework.

The SAT encoding is generated from a parsed and simplified representation of the input formula. Simplification reduces both the size of the encoding and the estimated size of countermodels, often yielding substantial performance improvements. The resulting SMT problem is then submitted to the Z3 solver. If a countermodel exists, Deo-SMT extracts a minimal one and presents it in several human-readable formats, including textual, matrix, and graph representations, facilitating inspection and explanation of failed entailments.

Experimental evaluation compared Deo-SMT with the Isabelle/HOL-based framework of Parent and Benz Müller [PB23] on a benchmark of 10,460 formulas. The SMT-based approach achieved an average speed-up of approximately a factor of $\times 31.8$. While the Isabelle/HOL approach is primarily aimed at meta-logical analysis and formal verification of logical properties, Deo-SMT focuses on efficient object-level reasoning and countermodel generation. In particular, its ability to compute minimal countermodels and present them in human-readable graphical, matrix, and textual formats makes it well suited for practical normative reasoning tasks, where countermodels can serve as explanations of why a given obligation does not follow or why a collection of norms exhibits unexpected behaviour.

4.8 GL-based hypersequent calculi for Åqvist's logics

In this section, we employ the finite-model semantics developed earlier in the chapter to obtain a uniform analytic proof theory for all four of Åqvist's logics. Our approach extends the **GL**-based methodology applied to **PCL** in Chapter 3. However, unlike

⁶<https://lego.guest.logic.at/>

$$\begin{array}{c}
 \frac{\begin{array}{c} \textcircled{O}(B|A), \Gamma \vdash \boxed{A}, \Delta \mid \mathcal{H} \quad \textcircled{O}(B|A), \Gamma \vdash \boxed{\neg A}, \Delta \mid \mathcal{H} \quad \boxed{B}, \textcircled{O}(B|A), \Gamma \vdash \Delta \mid \mathcal{H} \end{array}}{\textcircled{O}(B|A), \Gamma \vdash \Delta \mid \mathcal{H}} (\textcircled{O}L^\kappa) \\
 \\
 \frac{\boxed{A, \neg A \vdash B} \mid \Gamma \vdash \textcircled{O}(B|A), \Delta \mid \mathcal{H}}{\Gamma \vdash \textcircled{O}(B|A), \Delta \mid \mathcal{H}} (\textcircled{O}R^\kappa) \quad \frac{\textcircled{O}(B|A), \Gamma \vdash \Delta \mid \textcircled{O}(B|A), \Sigma \vdash \Pi \mid \mathcal{H}}{\textcircled{O}(B|A), \Gamma \vdash \Delta \mid \Sigma \vdash \Pi \mid \mathcal{H}} (com_{\textcircled{O}}^\kappa)
 \end{array}$$

Figure 4.3: Kleene-style hypersequent rules for conditional obligations from [COP22]. The addition in each premise w.r.t. the conclusion is highlighted.

PCL, all Åqvist logics contain **S5** as a fragment so instead of standard sequent calculi we will employ hypersequent calculi.

As discussed in Sec. 4.2, a number of proof calculi for Åqvist's logics have already been proposed, based on the idea of introducing the auxiliary modality $\Box$ into the language. In particular, Ciabattoni, Olivetti and Parent [COP22] developed a particularly natural hypersequent calculus for **E** based on the observation that conditional obligation $\textcircled{O}(B|A)$ can be represented as $\Box_{\mathbf{S5}}((A \wedge \Box \neg A) \rightarrow B)$, where $\Box$ is interpreted as a **K**-modality. Their calculus extends the hypersequent calculus $\mathcal{HK}_{\mathbf{S5}}$ for **S5** (Fig. 2.5) with the rules shown in Fig. 4.3, which decompose obligations according to this embedding and propagate them between hypersequent components. Additionally, there is a single elimination rule for the modality $\Box$, reflecting the modality elimination rule in **K** :

$$\frac{\Gamma^{\Box \downarrow} \vdash A \mid \Gamma \vdash \Box A, \Delta \mid \mathcal{H}}{\Gamma \vdash \Box A, \Delta \mid \mathcal{H}} (\Box_{\mathbf{E}}),$$

where $\Gamma^{\Box \downarrow} = \{A \mid \Box A \in \Gamma\}$.

Calculi employing $\Box$ -modalities have also been proposed for the stronger logics **F**, **F_{CM}**, and **G**. However, these calculi depart significantly from the natural bimodal treatment of **E**, resulting in proof systems that are neither particularly intuitive nor modular with respect to one another.

In this section, we introduce new calculi for **F**, **F_{CM}**, and **G**, which are modular and intuitive extensions of the calculus for **E** from [COP22]. We obtain them by modifying only the rules governing the $\Box$ -modality, while leaving the treatment of obligations unchanged. The introduced rules have clear semantic interpretations, reflecting the frame properties of the logic under consideration, which allows us to obtain direct and modular semantic proofs of soundness and completeness. This provides a uniform proof-theoretic account of the entire Åqvist family, directly connected to its finite-model semantics.

4.8.1 GL-based calculus for $\mathbf{F}_{\mathbf{CM}}$

The simplest case is that of $\mathbf{F}_{\mathbf{CM}}$, since its finite-model characterization—finite transitive preferential models—coincides directly with the semantics used for $\mathbf{PCL}$. Consequently, to obtain a calculus for $\mathbf{F}_{\mathbf{CM}}$, it suffices to replace the $\Box$ -elimination rule of $\mathcal{H}_{\mathbf{E}}$ with the corresponding **GL**-rule. Thus, the calculus $\mathcal{H}_{\mathbf{F}_{\mathbf{CM}}}$ extends the hypersequent calculus $\mathcal{HK}_{\mathbf{S5}}$ (Fig. 2.5,p.23) with the obligation rules from $\mathcal{H}_{\mathbf{E}}$ (Fig. 4.3) and the rule

$$\frac{\Gamma^{\Box\downarrow}, \Gamma^{\Box}, \Box A \vdash A \mid \Gamma \vdash \Box A, \Delta \mid \mathcal{H}}{\Gamma \vdash \Box A, \Delta \mid \mathcal{H}} (\Box_{\mathbf{F}_{\mathbf{CM}}}),$$

where $\Gamma^{\Box\downarrow} = \{A \mid \Box A \in \Gamma\}$ and $\Gamma^{\Box} = \{\Box A \mid \Box A \in \Gamma\}$.

Like $\mathcal{H}_{\mathbf{E}}$, the calculus $\mathcal{H}_{\mathbf{F}_{\mathbf{CM}}}$ enjoys a relaxed form of the subformula property: every derivation starting from a hypersequent $\Gamma_1 \vdash \Delta_1 \mid \dots \mid \Gamma_n \vdash \Delta_n$ contains only formulas from $\bigcup_{1 \leq i \leq n} (\Gamma_i \cup \Delta_i \cup \{\Box \neg A \mid \bigcirc(B|A) \in \Gamma_i \cup \Delta_i\})$. Since only finitely many such formulas can occur, there are only finitely many possible hypersequent components. Consequently, for every root hypersequent there is a finite bound on the length of the branches in the derivations, as long as there are no *redundant rule applications*.

Definition 25. A rule application is *redundant* if it has a premise such that, for every component $\Gamma \vdash \Delta$ occurring in it, there exists a component $\Gamma' \vdash \Delta'$ in the conclusion with all formulas in Γ occurring in Γ' and all formulas in Δ occurring in Δ' .

Such rule applications are indeed redundant, since they do not introduce any new information in the selected premise and any derivation of the this premise can equally well be carried out from the conclusion. We call a hypersequent *saturated* if no non-redundant rule application to it is possible. A finite transitive preferential countermodal can be extracted from a hypersequent saturated with respect to the rules of $\mathcal{H}_{\mathbf{F}_{\mathbf{CM}}}$, which provides a direct semantic completeness proof for the calculus.

Theorem 17. Hypersequent $\Gamma_1 \vdash \Delta_1 \mid \dots \mid \Gamma_n \vdash \Delta_n$ is derivable in $\mathcal{H}_{\mathbf{F}_{\mathbf{CM}}}$ iff formula $\bigvee_{1 \leq i \leq n} \Box((\bigwedge_{A \in \Gamma_i} A) \rightarrow (\bigvee_{B \in \Delta_i} B))$ is valid in $\mathbf{F}_{\mathbf{CM}}$.

Proof. Soundness: Suppose towards contradiction that hypersequent $\mathcal{H}$ is derivable but there is some model M falsifying this hypersequent, i.e for any $i \in \{1, \dots, n\}$ there is a world w in this model such that $M, w \models A$ for all $A \in \Gamma_i$ and $M, w \not\models B$ for all $B \in \Delta_i$. It is straightforward to check for each rule of $\mathcal{H}_{\mathbf{F}_{\mathbf{CM}}}$, that if a model falsifies the conclusion of this rule then the same model falsifies at least one of the premises. Consider the rule $(\Box_{\mathbf{F}_{\mathbf{CM}}})$ for example. If there is a world w such that $M, w \models X$ for all $X \in \Gamma$ and $M, w \not\models \Box A$ then there should be a world u such that $u \succ w$ and $M, u \not\models A$. Due to smoothness there is a world v such that $v \in \max_{\succ}(\|\neg A\|^M)$ and either $v = u$ or $v \succ u$; in either case $v \succ w$. For every $\Box X \in \Gamma$, $M, w \models \Box X$ implies $M, v \models X$ and $M, v \models \Box X$ due to transitivity, also $M, v \models \Box A$ due to $v \in \max_{\succ}(\|\neg A\|^M)$, therefore the additional component in the premise of $(\Box_{\mathbf{F}_{\mathbf{CM}}})$ is falsified in v . Propagating this

reasoning step inductively through the derivation of $\mathcal{H}$ bottom-up, we arrive to the fact that the conclusion of the concluding rule is falsified in some model, which leads to an obvious contradiction.

Completeness: Suppose that hypersequent $\mathcal{H}$ is not derivable in $\mathcal{H}_{\mathbf{FCM}}$. Let us perform arbitrary non-redundant applications of rules of $\mathcal{H}_{\mathbf{FCM}}$ to it bottom-up until every branch is either closed or ends with a saturated hypersequent (which will happen after finite number of such applications). Due to non-derivability there should be at least one branch ending with a saturated sequent $\mathcal{H}'$ that can not be closed by a concluding rule. We will now construct a finite transitive model falsifying $\mathcal{H}'$, which is sufficient for completeness, since it will also falsify $\mathcal{H}$ due to the Kleene form of the rules. As the set of the worlds W we take the set of components in $\mathcal{H}'$. Valuation on the component $\Gamma \vdash \Delta$ is defined as $\Gamma \cap \text{Atm}$. Preference relation $\succeq$ on components is defined as follows: $(\Sigma \vdash \Pi) \succeq (\Gamma \vdash \Delta)$ when $\Gamma^{\Box} \cup \Gamma^{\Box\downarrow} \subseteq \Sigma$. It is easy to check that this preference relation is transitive. We thus get a finite transitive preferential model. Let us now establish the *truth lemma* for our countermodel construction: every component $\Gamma \vdash \Delta$ in $\mathcal{H}'$ satisfies all formulas in Γ and falsifies all formulas in Δ in the constructed model. We show it by induction on the size of a formula, measured as number of symbols, considering all propositional atoms, Boolean connectives, and unary modalities as one symbol, and $\bigcirc(B|A)$ as four symbols in addition to the sizes of A and B . For propositional atoms the truth lemma follows from the choice of valuation and the fact that the concluding rule can not be applied to $\mathcal{H}'$ (so atoms appearing on the right can not appear on the left). For formulas with Boolean connectives on top the truth lemma follows from saturation w.r.t. logical rules directly. Now consider the case of formulas with $\Box$ on top. If $\Box A \in \Gamma$ then due to saturation w.r.t. $(\text{com}_{\mathbf{S5}}^K)$ formula $\Box A$ is present on the left in any component of $\mathcal{H}'$, then due to saturation w.r.t. $(\Box_{\mathbf{S5}} L^K)$ formula A is present on the left in any component, which means that A is satisfied in all worlds of the constructed model by inductive hypothesis. If $\Box A \in \Delta$ then due to saturation w.r.t. $(\Box_{\mathbf{S5}} R^K)$ there is a component in $\mathcal{H}'$ with A on the right, which implies by inductive hypothesis that A is not true in all worlds in the constructed model. Now consider the case of formulas with $\Box$ on top. If $\Box A \in \Gamma$, consider any component $\Sigma \vdash \Pi$ strictly more preferred than $\Gamma \vdash \Delta$, by definition of the preference relation $A \in \Sigma$, which means A is satisfied by component $\Sigma \vdash \Pi$ in the model, thus $\Box A$ is satisfied by $\Gamma \vdash \Delta$. If $\Box A \in \Delta$, then due to saturation w.r.t. $(\Box_{\mathbf{FCM}})$ there is a component $\Sigma \vdash \Pi$ in $\mathcal{H}'$ such that $\Gamma^{\Box\downarrow} \cup \Gamma^{\Box} \cup \{\Box A\} \subseteq \Sigma$ and $A \in \Pi$, then $\Sigma \vdash \Pi$ is strictly preferable to $\Gamma \vdash \Delta$ (note that $\Gamma \vdash \Delta$ can not be preferable to $\Sigma \vdash \Pi$ since Σ contains $\Box A$ and Γ does not, otherwise the concluding rule could be applied to $\mathcal{H}'$), while A is false in $\Sigma \vdash \Pi$, so $\Box A$ is false in $\Gamma \vdash \Delta$. Finally, consider the case of formulas with $\bigcirc$ on top. If $\bigcirc(B|A) \in \Gamma$ then due to saturation w.r.t. $(\text{com}_{\bigcirc}^K)$ formula $\bigcirc(B|A)$ is present on the left in any component of $\mathcal{H}'$, then due to saturation w.r.t. $(\bigcirc L^K)$ each component either contains B on the left, or A on the right, or $\Box \neg A$ on the right, which means by inductive hypothesis that each component either satisfies B or is not $\succ$ -maximal world in the truth set of A (note that $\Box \neg A$ is strictly smaller than $\bigcirc(A|B)$ w.r.t. our definition of formula size). If $\bigcirc(B|A) \in \Delta$ then due to saturation w.r.t. $(\bigcirc R^K)$ there is a component in $\mathcal{H}'$ with A and $\Box \neg A$ on the left and B on the

right, which implies by inductive hypothesis that this component is $\succ$ -maximal world in the truth set of A and B is not true in it, so $\bigcirc(B|A)$ is not true in the constructed model. We thus showed that the constructed model falsifies the saturated hypersequent $\mathcal{H}'$, which concludes the proof of completeness. $\square$

Example 8. Consider again the instance of (RM)

$$\bigcirc(c | a) \vdash \bigcirc(\neg b | a), \bigcirc(c | a \wedge b).$$

which is not valid in $\mathbf{FCM}$. Let us perform bottom-up proof search in $\mathcal{H}_{\mathbf{FCM}}$ to reconstruct a countermodel from a failed derivation.

We show an open branch, which ends with a saturated hypersequent consisting of the following four components:

$$\begin{aligned} \Gamma_1 &= \bigcirc(c | a), c \vdash \bigcirc(\neg b | a), \bigcirc(c | a \wedge b), \\ \Gamma_2 &= \bigcirc(c | a), a, b, c, \boxminus \neg a \vdash \neg b, \\ \Gamma_3 &= \bigcirc(c | a), a \wedge b, a, b, \boxminus \neg(a \wedge b) \vdash c, \boxminus \neg a, \\ \Gamma_4 &= \bigcirc(c | a), a, c, \neg(a \wedge b), \boxminus \neg(a \wedge b), \boxminus \neg a \vdash \neg a, a \wedge b, b. \end{aligned}$$

In the branch *bell*, once a component reaches its final form, we abbreviate it by the corresponding Γ_i . Applications of the rule of **LK** are collected into single steps.

$$\begin{array}{c} \frac{\bigcirc(c | a), a, \boxed{c}, \neg(a \wedge b), \boxminus \neg(a \wedge b), \boxminus \neg a \vdash \neg a, a \wedge b, b \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(c | a), \boxed{a}, \neg(a \wedge b), \boxminus \neg(a \wedge b), \boxminus \neg a \vdash \neg a, \boxed{a \wedge b, b} \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1} (\bigcirc L^\mathcal{K}) \\ \hline \frac{\bigcirc(c | a), \neg(a \wedge b), \boxminus \neg(a \wedge b), \boxminus \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1}{\neg(a \wedge b), \boxminus \neg(a \wedge b), \boxminus \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1} (com_{\bigcirc}^\mathcal{K}) \\ \hline \frac{\neg(a \wedge b), \boxminus \neg(a \wedge b), \boxminus \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(c | a), a \wedge b, a, b, \boxminus \neg(a \wedge b) \vdash c, \boxminus \neg a \mid \Gamma_2 \mid \Gamma_1} (\boxminus \mathbf{FCM}) \\ \hline \frac{\bigcirc(c | a), a \wedge b, a, b, \boxminus \neg(a \wedge b) \vdash c, \boxminus \neg a \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(c | a), a \wedge b, \boxed{a, b}, \boxminus \neg(a \wedge b) \vdash c \mid \Gamma_2 \mid \Gamma_1} (\bigcirc L^\mathcal{K}) \\ \hline \frac{\bigcirc(c | a), a \wedge b, \boxed{a, b}, \boxminus \neg(a \wedge b) \vdash c \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(c | a), a \wedge b, \boxminus \neg(a \wedge b) \vdash c \mid \Gamma_2 \mid \Gamma_1} LK \\ \hline \frac{\bigcirc(c | a), a \wedge b, \boxminus \neg(a \wedge b) \vdash c \mid \Gamma_2 \mid \Gamma_1}{a \wedge b, \boxminus \neg(a \wedge b) \vdash c \mid \Gamma_2 \mid \Gamma_1} (com_{\bigcirc}^\mathcal{K}) \\ \hline \frac{a \wedge b, \boxminus \neg(a \wedge b) \vdash c \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(c | a), a, b, \boxed{c}, \boxminus \neg a \vdash \neg b \mid \Gamma_1} (\bigcirc R^\mathcal{K}) \\ \hline \frac{\bigcirc(c | a), a, b, \boxed{c}, \boxminus \neg a \vdash \neg b \mid \Gamma_1}{\bigcirc(c | a), a, \boxed{b}, \boxminus \neg a \vdash \neg b \mid \Gamma_1} (\bigcirc L^\mathcal{K}) \\ \hline \frac{\bigcirc(c | a), a, \boxed{b}, \boxminus \neg a \vdash \neg b \mid \Gamma_1}{\bigcirc(c | a), a, \boxminus \neg a \vdash \neg b \mid \Gamma_1} LK \\ \hline \frac{\bigcirc(c | a), a, \boxminus \neg a \vdash \neg b \mid \Gamma_1}{a, \boxminus \neg a \vdash \neg b \mid \Gamma_1} (com_{\bigcirc}^\mathcal{K}) \\ \hline \frac{a, \boxminus \neg a \vdash \neg b \mid \Gamma_1}{\bigcirc(c | a), \boxed{c} \vdash \bigcirc(\neg b | a), \bigcirc(c | a \wedge b)} (\bigcirc R^\mathcal{K}) \\ \hline \frac{\bigcirc(c | a), \boxed{c} \vdash \bigcirc(\neg b | a), \bigcirc(c | a \wedge b)}{\bigcirc(c | a) \vdash \bigcirc(\neg b | a), \bigcirc(c | a \wedge b)} (\bigcirc L^\mathcal{K}) \end{array}$$

Taking the four components of the saturated hypersequent as worlds, the valuation is determined by the propositional atoms on their left:

$$V(\Gamma_1) = \{c\}, \quad V(\Gamma_2) = \{a, b, c\}, \quad V(\Gamma_3) = \{a, b\}, \quad V(\Gamma_4) = \{a, c\}.$$

And the following preference relation is reconstructed:

$$\Gamma_1 \preceq \Gamma_i \text{ for } i \in \{1, \dots, 4\} \quad \text{and} \quad \Gamma_3 \preceq \Gamma_4.$$

The best a -worlds are Γ_2 and Γ_4 , so $\bigcirc(c \mid a)$ holds while $\bigcirc(\neg b \mid a)$ does not. The best $a \wedge b$ -worlds are Γ_2 and Γ_3 , and $\Gamma_3 \not\models c$, so $\bigcirc(c \mid a \wedge b)$ is false. Hence the reconstructed model falsifies the initial hypersequent.

4.8.2 GL-based calculus for $\mathbf{G}$

The logic $\mathbf{G}$ differs from $\mathbf{F}_{\mathbf{CM}}$ by imposing the additional requirement that the preference relation be total. Proof-theoretically, this property can be captured by the following communication rule:

$$\frac{\Sigma^{\Box\downarrow}, \Sigma^{\Box}, \Gamma \vdash \Delta \mid \Sigma \vdash \Pi \mid \mathcal{H} \quad \Gamma \vdash \Delta \mid \Gamma^{\Box\downarrow}, \Gamma^{\Box}, \Sigma \vdash \Pi \mid \mathcal{H} \quad \Sigma^{\Box}, \Gamma \vdash \Delta \mid \Gamma^{\Box}, \Sigma \vdash \Pi \mid \mathcal{H}}{\Gamma \vdash \Delta \mid \Sigma \vdash \Pi \mid \mathcal{H}} \text{ (com}_{\Box}\text{)}$$

Intuitively, the two components involved in the rule represent two worlds of a preferential model. By totality, either one of these worlds is strictly preferred to the other, in which case all formulas under $\Box$ -modalities from the less preferred world can be propagated to the more preferred one, or the worlds are mutually preferred, in which case $\Box$ -formulas can be exchanged between them. Thus, the rule enforces precisely the interaction between worlds induced by totality of the preference relation. We therefore define the calculus $\mathcal{H}_{\mathbf{G}}$ by extending $\mathcal{H}_{\mathbf{F}_{\mathbf{CM}}}$ with the rule $(\text{com}_{\Box})$. The proof of the soundness and completeness can proceed in the same way as for $\mathcal{H}_{\mathbf{F}_{\mathbf{CM}}}$, we only need to verify that the additional rule is sound w.r.t. total preference relation and that saturation w.r.t. this rule ensures totality of the preference relation in the extracted countermodel.

Soundness of the resulting calculus can be verified directly with respect to the semantics of $\mathbf{G}$.

Theorem 18. *If hypersequent $\Gamma_1 \vdash \Delta_1 \mid \dots \mid \Gamma_1 \vdash \Delta_n$ is derivable in $\mathcal{H}_{\mathbf{G}}$ then formula $\bigvee_{1 \leq i \leq n} \Box((\bigwedge_{A \in \Gamma_i} A) \rightarrow (\bigwedge_{B \in \Delta_i} B))$ is valid in $\mathbf{G}$.*

Proof. Analogous to the proof of soundness of $\mathcal{H}_{\mathbf{F}_{\mathbf{CM}}}$ (Th. 17). All rules of $\mathcal{H}_{\mathbf{F}_{\mathbf{CM}}}$ are sound w.r.t. models of $\mathbf{G}$ as they were shown sound w.r.t. models of $\mathbf{F}_{\mathbf{CM}}$ already. Soundness of $(\text{com}_{\Box})$ relies on totality of the preference relation. Specifically, suppose the conclusion of $(\text{com}_{\Box})$ is falsified in some model M , then there are worlds w_1 and w_2 in M such that w_1 satisfies all formulas in Γ and falsifies all formulas in Δ , while

w_2 satisfies all formulas in Σ and falsifies all formulas in Π . Due to the totality of the preference relation either $w_1 \succ w_2$ or $w_2 \succ w_1$ or both $w_1 \succeq w_2$ and $w_2 \succeq w_1$. In the first case, w_1 satisfies all formulas in $\Sigma^{\Box\downarrow}$ and also all formulas in $\Sigma^{\Box}$ due to transitivity of the preference relation, so M falsifies the first premise. The second case is symmetrical, and in this case M falsifies the second premise. In the third case, $z \succ w_1$ is equivalent to $z \succ w_2$ for every world z due to transitivity of the preference relation, so w_1 satisfies all formulas in $\Sigma^{\Box}$ and w_2 satisfies all formulas in $\Gamma^{\Box}$, so M falsifies the third premise. $\square$

For completeness, it is convenient to replace totality by the following *modularity* condition on the strict version of the preference relation, as suggested in [LM92].

Definition 26. *Preference relation $\succeq$ is called modular if for all $x, y, z \in W$, if $x \succ y$ then either $x \succ z$ or $z \succ y$.*

Modularity together with transitivity and finitedness provides an equivalent characterization of $\mathbf{G}$.

Lemma 14. *A formula is valid in $\mathbf{G}$ iff it is valid in all finite preferential models, in which preference relation $\succeq$ is transitive and modular.*

Proof. Right-to-left direction follows from the characterization of $\mathbf{G}$ via finite models with the transitive and total preference relation $\succeq$, since for every such model $\succ$ is modular: by totality $x \not\succ z$ implies $z \succeq x$ and $z \not\succ y$ implies $y \succeq z$, which due to transitivity would contradict $x \succ y$. For left-to-right direction, suppose $M = \langle W, \succeq, V \rangle$ is a finite preferential model, $\succeq$ is transitive and modular and $M, w \not\models A$ for some $w \in W$. We will show that A is then falsified by some transitive and total model too. Namely, consider the model $M' = \langle W, \succeq', V \rangle$ where $x \succeq' y$ iff $x \not\succ y$. Relation $\succeq'$ is transitive: $x \not\succ y$, $y \not\succ z$, and $x \prec z$ together would contradict the modularity of $\succeq$. It is also trivially reflexive and total, so M' is a $\mathbf{G}$ -model. Also $\succ' = \succ$, so we can show by trivial induction that $M, w \not\models A$ is equivalent to $M', w \not\models A$, concluding the proof. $\square$

This characterization allows the completeness proof to proceed in the same way as for $\mathbf{F}_{\mathbf{CM}}$. It suffices to show that the countermodel extracted from a hypersequent saturated with respect to the rules of $\mathcal{H}_{\mathbf{G}}$ satisfies modularity.

Theorem 19. *If formula $\bigvee_{1 \leq i \leq n} \Box((\bigwedge_{A \in \Gamma_i} A) \rightarrow (\bigwedge_{B \in \Delta_i} B))$ is valid in $\mathbf{G}$ then hypersequent $\Gamma_1 \vdash \Delta_1 \mid \dots \mid \Gamma_1 \vdash \Delta_n$ is derivable in $\mathcal{H}_{\mathbf{G}}$.*

Proof. Analogous to the proof of completeness of $\mathcal{H}_{\mathbf{F}_{\mathbf{CM}}}$ (Th. 17). We take the same countermodel reconstructed from the saturated hypersequent. Since $\mathcal{H}_{\mathbf{G}}$ contains all the rules of $\mathcal{H}_{\mathbf{F}_{\mathbf{CM}}}$ the reasoning in that proof also goes for $\mathcal{H}_{\mathbf{G}}$, therefore the reconstructed finite and transitive countermodel indeed falsifies the saturated hypersequent. It remains to show that in the case of $\mathcal{H}_{\mathbf{G}}$ the preference relation in that countermodel is modular. Suppose some component $\Sigma \vdash \Pi$ is strictly preferred to some component $\Gamma \vdash \Delta$ and consider some third component $\Theta \vdash \Xi$. Let us consider two cases:

- $\Theta^{\Box} \subseteq \Gamma^{\Box}$. Then $\Sigma \vdash \Pi$ is preferred also to $\Theta \vdash \Xi$ (since in such case $\Theta^{\Box\downarrow} \subseteq \Gamma^{\Box\downarrow}$). So either $\Sigma \vdash \Pi$ is strictly preferred to $\Theta \vdash \Xi$ or $\Sigma \vdash \Pi$ and $\Theta \vdash \Xi$ are mutually preferred, the latter implying that $\Theta \vdash \Xi$ is strictly preferred to $\Gamma \vdash \Delta$ due to transitivity, thus satisfying the modularity requirement.
- $\Theta^{\Box} \not\subseteq \Gamma^{\Box}$. Then we can apply the saturation condition for rule $(com_{\Box}^{\mathcal{K}})$ to components $\Gamma \vdash \Delta$ and $\Theta \vdash \Xi$, and conclude that in this case $\Gamma^{\Box} \cup \Gamma^{\Box\downarrow} \subseteq \Theta$. Therefore $\Theta \vdash \Xi$ is preferred to $\Gamma \vdash \Delta$. So either $\Theta \vdash \Xi$ is strictly preferred to $\Gamma \vdash \Delta$ or $\Theta \vdash \Xi$ and $\Gamma \vdash \Delta$ are mutually preferred, the latter implying that $\Sigma \vdash \Pi$ is strictly preferred to $\Theta \vdash \Xi$, thus satisfying the modularity requirement. $\square$

Example 9. Consider the conditional excluded middle

$$\vdash \bigcirc(b \mid a), \bigcirc(\neg b \mid a),$$

which is not valid in $\mathbf{G}$. Let us perform bottom-up proof search in $\mathcal{H}_{\mathbf{G}}$ to reconstruct a countermodel from a failed derivation.

The following open branch ends with a saturated hypersequent:

$$\frac{\frac{\frac{a, \boxed{b}, \Box \neg a \vdash \neg b \mid a, \Box \neg a \vdash b \mid \vdash \bigcirc(b \mid a), \bigcirc(\neg b \mid a)}{a, \Box \neg a \vdash \neg b \mid a, \Box \neg a \vdash b \mid \vdash \bigcirc(b \mid a), \bigcirc(\neg b \mid a)} LK}{a, \Box \neg a \vdash \neg b \mid a, \Box \neg a \vdash b \mid \vdash \bigcirc(b \mid a), \bigcirc(\neg b \mid a)} (\bigcirc R^{\mathcal{K}})}{\frac{a, \Box \neg a \vdash b \mid \vdash \bigcirc(b \mid a), \bigcirc(\neg b \mid a)}{\vdash \bigcirc(b \mid a), \bigcirc(\neg b \mid a)} (\bigcirc R^{\mathcal{K}})}$$

The resulting hypersequent is saturated also with respect to the additional communication rule of $\mathcal{H}_{\mathbf{G}}$: the two newly introduced components have the same $\Box$ -context, while every application involving the initial component has a redundant premise.

Taking the three components of the saturated hypersequent as worlds w_1, w_2, w_3 , from right to left, the reconstructed valuation is

$$V(w_1) = \emptyset, \quad V(w_2) = \{a\}, \quad V(w_3) = \{a, b\}.$$

The reconstructed strict preference relation is

$$w_1 \prec w_2, \quad w_1 \prec w_3,$$

with no strict comparison between w_2 and w_3 . It is transitive and modular, and hence characterizes a $\mathbf{G}$ -countermodel by Lem. 14.

The best a -worlds are w_2 and w_3 . Since they disagree on b , both $\bigcirc(b \mid a)$ and $\bigcirc(\neg b \mid a)$ are false. Hence the reconstructed model falsifies the initial hypersequent.

4.8.3 GL-based calculus for $\mathbf{F}$

The limitedness condition underlying $\mathbf{F}$ is quite challenging to be captured via sequent rules, as demonstrated in [COP⁺23]. It is also not clear how to directly capture our alternative finite-model characterization via acyclicity of the preference relation by means of inference rules. However, we can exploit our LEGO model construction for $\mathbf{F}$ to derive an alternative characterization that is more suitable for sequent rule treatment and can in fact be captured by adapting the $\mathbf{GL}$ rule.

For this, we introduce a weaker form of transitivity, which we call *tail-transitivity*. Intuitively, transitivity is required only for triples $x \succ y \succ z$ that occur as the tail of a longer preference chain $x \succ y \succ z \succ w$. Formally, we have the following definition.

Definition 27. A preference relation $\succ$ is called *tail-transitive* if for any worlds $w, x, y, z \in W$, $x \succ y \succ z \succ w$ implies $x \succ z$.

We can use Lem. 13 to establish that finitedness tail-transitivity provides an equivalent semantic characterization of $\mathbf{F}$.

Lemma 15. A formula is valid in $\mathbf{F}$ iff it is valid in all finite preferential models whose preference relation is tail-transitive.

Proof. We apply Lem. 13 to derive this new characterization from the standard characterization via limitedness. It is easy to check that $SMC^{\mathbf{F}}(A, M)$ (Fig. 4.2) satisfies tail-transitivity. Let us now show that a finite and tail-transitive model satisfies limitedness. Suppose there is a world w_0 in M satisfying A , consider (any) longest path $w_0 \prec w_1 \prec w_2 \prec \dots$ of worlds from $\|A\|^M$ starting from w_0 . No world can occur on this path more than once, since otherwise the cycle could be repeated after the second occurrence and tail-transitivity could be applied to this cycle leading to the contradictory statement $x \succ x$ for some world x . Since W is finite and there can be no repetitions on the path, the path is finite and the last world on the path belongs to $\max_{\succ}(\|A\|^M)$ by definition, establishing limitedness of the model. $\square$

To capture tail-transitivity proof-theoretically, we further extend the language with a constant $\blacktriangle$, interpreted as the existence of a strictly less preferred world:

$$M, w \models \blacktriangle \quad \text{when} \quad \exists u \in W, w \succ u.$$

We can then use this constant to control applications of the $\mathbf{GL}$ rule, which only holds in the part of the model satisfying transitivity. Accordingly, our calculus $\mathcal{H}_{\mathbf{F}}$ for $\mathbf{F}$ extends $\mathcal{HK}_{\mathbf{S5}}$ and the obligation rules of $\mathcal{H}_{\mathbf{E}}$ with two elimination rules for the modality $\Box$:

$$\frac{\blacktriangle, \Gamma^{\Box\downarrow} \vdash A \mid \Gamma \vdash \Box A, \Delta \mid \mathcal{H}}{\Gamma \vdash \Box A, \Delta \mid \mathcal{H}} (\Box_{\mathbf{F}}) \quad \frac{\blacktriangle, \Gamma^{\Box\downarrow}, \Gamma^{\Box}, \Box A \vdash A \mid \blacktriangle, \Gamma \vdash \Box A, \Delta \mid \mathcal{H}}{\blacktriangle, \Gamma \vdash \Box A, \Delta \mid \mathcal{H}} (\Box_{\mathbf{F}}^{\blacktriangle})$$

where $\Gamma^{\Box\downarrow} = \{A \mid \Box A \in \Gamma\}$ and $\Gamma^{\Box} = \{\Box A \mid \Box A \in \Gamma\}$. The first rule corresponds to the elimination rule in **K**, moving to a more preferred world and introducing $\blacktriangle$ into the premise. The second is a **GL** rule whose application requires $\blacktriangle$ to be present in the conclusion.

Soundness and completeness are established by the same semantic methodology as for **F_{CM}** and **G**. The constant $\blacktriangle$ is incorporated into the definition of the preference relation reconstructed from saturated hypersequents, ensuring that the resulting countermodels satisfy tail-transitivity.

Theorem 20. *Hypersequent $\Gamma_1 \vdash \Delta_1 \mid \dots \mid \Gamma_n \vdash \Delta_n$ is derivable in $\mathcal{H}_F$ iff formula $\bigvee_{1 \leq i \leq n} \Box((\bigwedge_{A \in \Gamma_i} A) \rightarrow (\bigwedge_{B \in \Delta_i} B))$ is valid in **F**.*

Proof. Analogous to the proof of Th. 17.

Soundness of all rules apart from $(\Box_F)$ and $(\Box_F^\blacktriangle)$ is the same as in the case of **F_{CM}**. Soundness of $(\Box_F)$ follows from the fact that if there is a world falsifying $\Box A$ and satisfying formulas in Γ then there is a world strictly preferred to it, falsifying A and satisfying formulas in $\Gamma^{\Box\downarrow}$, and this world also satisfies $\blacktriangle$ by definition. Soundness of $(\Box_F^\blacktriangle)$ follows from tail-transitivity. Specifically, suppose some world z satisfies $\blacktriangle$ and formulas in Γ and falsifies $\Box A$, then there is a world w such that $z \succ w$, and also a world y such that $y \succ z$ and $M, y \not\models A$. Similarly to the proof of Lem. 15 we can consider the longest increasing path of worlds in $\|\neg A\|^M$ starting from y , it will end with some world x_m such that $x_m \in \max_{\succ}(\|\neg A\|^M)$ and $x_m \succ \dots \succ x_1 \succ y \succ z \succ w$. Applying tail-transitivity to this chain of worlds, we get $x_m \succ z$, which implies that x_m satisfies all formulas in $\Gamma^{\Box\downarrow}$. x_m also satisfies all formulas in $\Gamma^{\Box}$ since $u \succ x_m$ implies $u \succ z$ by tail-transitivity. x_m also satisfies $\Box A$ and falsifies A since $x_m \in \max_{\succ}(\|\neg A\|^M)$. Thus, the additional component in the premise of $(\Box_F^\blacktriangle)$ is falsified in x_m .

For completeness, we can consider a similar reconstruction of a countermodel from a saturated hypersequent as in the proof of Th. 17, with components as worlds, but we need to change the preference relation to accommodate the presence of constant $\blacktriangle$. Namely, in the modified definition, $(\Sigma \vdash \Pi) \succeq (\Gamma \vdash \Delta)$ when $\{\blacktriangle\} \cup \Gamma^{\Box\downarrow} \subseteq \Sigma$ and either $\Gamma^{\Box} \subseteq \Sigma$ or $\blacktriangle \notin \Gamma$. This relation is tail-transitive: if $(\Gamma_1 \vdash \Delta_1) \succ (\Gamma_2 \vdash \Delta_2) \succ (\Gamma_3 \vdash \Delta_3) \succ (\Gamma_4 \vdash \Delta_4)$ then by definition $\blacktriangle \in \Gamma_i$ for any $i \in \{1, 2, 3\}$, so $\Gamma_3^{\Box} \subseteq \Gamma_2^{\Box} \subseteq \Gamma_1^{\Box} \subseteq \Gamma_1$ and $\Gamma_3^{\Box\downarrow} \subseteq \Gamma_2^{\Box\downarrow} \subseteq \Gamma_1$, implying $(\Gamma_1 \vdash \Delta_1) \succeq (\Gamma_3 \vdash \Delta_3)$, while also $(\Gamma_3 \vdash \Delta_3) \not\succeq (\Gamma_1 \vdash \Delta_1)$ otherwise $\Gamma_2^{\Box} \subseteq \Gamma_1^{\Box} \subseteq \Gamma_3$ and $\Gamma_2^{\Box\downarrow} \subseteq \Gamma_1^{\Box\downarrow} \subseteq \Gamma_3$ contradicting $(\Gamma_2 \vdash \Delta_2) \succ (\Gamma_3 \vdash \Delta_3)$, so $(\Gamma_1 \vdash \Delta_1) \succ (\Gamma_3 \vdash \Delta_3)$ as required. With this definition we can prove that the constructed model falsifies the saturated hypersequent the same way as in Th. 17, only the case of $\Box$ -formulas changes. Suppose that $\Box A \in \Gamma$ for some component $\Gamma \vdash \Delta$ in the saturated hypersequent, then for any component $\Sigma \vdash \Pi$ strictly preferred to it, $A \in \Sigma$ by definition of preference relation, so A is satisfied in $\Sigma \vdash \Pi$, so $\Box A$ is satisfied in $\Gamma \vdash \Delta$. Now suppose that $\Box A \in \Delta$ for some component $\Gamma \vdash \Delta$ in the saturated hypersequent. We need to consider two cases. If $\blacktriangle \notin \Gamma$ then due to saturation w.r.t. $(\Box_F)$ there is a component $\Sigma \vdash \Pi$ such that $\Gamma^{\Box\downarrow} \cup \{\blacktriangle\} \subseteq \Sigma$ and $A \in \Pi$, this component is strictly

preferred to $\Gamma \vdash \Delta$ by definition of the preference relation and falsifies A by inductive hypothesis, so in this case $\Box A$ is not true in $\Gamma \vdash \Delta$ as required. If $\blacktriangle \in \Gamma$ then due to saturation w.r.t. $(\Box_{\mathbf{F}}^{\blacktriangle})$ there is a component $\Sigma \vdash \Pi$ such that $\Gamma^{\Box} \cup \Gamma^{\Box\downarrow} \cup \{\blacktriangle, \Box A\} \subseteq \Sigma$ and $A \in \Pi$, this component is strictly preferred to $\Gamma \vdash \Delta$ since $\Box A \in \Sigma^{\Box} \setminus \Gamma^{\Box}$ and it falsifies A by inductive hypothesis, so in this case $\Box A$ is not true in $\Gamma \vdash \Delta$ as required too. This concludes the completeness proof. $\square$

Example 10. Consider the instance of (CM) $\bigcirc(b \mid a), \bigcirc(c \mid a) \vdash \bigcirc(c \mid a \wedge b)$, which is not valid in $\mathbf{F}$. Let us perform bottom-up proof search in $\mathcal{H}_{\mathbf{F}}$ to reconstruct a countermodel from a failed derivation.

$$\begin{array}{c}
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, a, b, \boxed{c}, \Box \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, a, \boxed{b}, \Box \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1} (\bigcirc L^{\mathcal{K}}) \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, a, \boxed{b}, \Box \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, \boxed{a}, \Box \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1} (\bigcirc L^{\mathcal{K}}) \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, \boxed{a}, \Box \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, \Box \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1} LK \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, \Box \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(b \mid a), \blacktriangle, \Box \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1} (com_{\bigcirc}^{\mathcal{K}}) \\
\frac{\bigcirc(b \mid a), \blacktriangle, \Box \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1}{\blacktriangle, \Box \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1} (com_{\bigcirc}^{\mathcal{K}}) \\
\frac{\blacktriangle, \Box \neg a \vdash \neg a \mid \Gamma_3 \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, a, \boxed{c}, \neg(a \wedge b) \vdash \neg a, a \wedge b, b, \Box \neg a \mid \Gamma_2 \mid \Gamma_1} (\Box_{\mathbf{F}}^{\blacktriangle}) \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, a, \boxed{c}, \neg(a \wedge b) \vdash \neg a, a \wedge b, b, \Box \neg a \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, a, \neg(a \wedge b) \vdash \neg a, a \wedge b, b, \Box \neg a \mid \Gamma_2 \mid \Gamma_1} (\bigcirc L^{\mathcal{K}}) \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, a, \neg(a \wedge b) \vdash \neg a, a \wedge b, b, \Box \neg a \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, \boxed{a}, \neg(a \wedge b) \vdash \neg a, \boxed{a \wedge b}, b \mid \Gamma_2 \mid \Gamma_1} (\bigcirc L^{\mathcal{K}}) \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, \boxed{a}, \neg(a \wedge b) \vdash \neg a, \boxed{a \wedge b}, b \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, \neg(a \wedge b) \vdash \neg a \mid \Gamma_2 \mid \Gamma_1} LK \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, \neg(a \wedge b) \vdash \neg a \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(b \mid a), \blacktriangle, \neg(a \wedge b) \vdash \neg a \mid \Gamma_2 \mid \Gamma_1} (com_{\bigcirc}^{\mathcal{K}}) \\
\frac{\bigcirc(b \mid a), \blacktriangle, \neg(a \wedge b) \vdash \neg a \mid \Gamma_2 \mid \Gamma_1}{\blacktriangle, \neg(a \wedge b) \vdash \neg a \mid \Gamma_2 \mid \Gamma_1} (com_{\bigcirc}^{\mathcal{K}}) \\
\frac{\blacktriangle, \neg(a \wedge b) \vdash \neg a \mid \Gamma_2 \mid \Gamma_1}{\bigcirc(b \mid a), \bigcirc(c \mid a), a \wedge b, a, b, \Box \neg(a \wedge b) \vdash c, \Box \neg a \mid \Gamma_1} (\Box_{\mathbf{F}}) \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), a \wedge b, a, b, \Box \neg(a \wedge b) \vdash c, \Box \neg a \mid \Gamma_1}{\bigcirc(b \mid a), \bigcirc(c \mid a), a \wedge b, a, b, \Box \neg(a \wedge b) \vdash c \mid \Gamma_1} (\bigcirc L^{\mathcal{K}}) \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), a \wedge b, a, b, \Box \neg(a \wedge b) \vdash c \mid \Gamma_1}{\bigcirc(b \mid a), \bigcirc(c \mid a), a \wedge b, \Box \neg(a \wedge b) \vdash c \mid \Gamma_1} LK \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), a \wedge b, \Box \neg(a \wedge b) \vdash c \mid \Gamma_1}{\bigcirc(b \mid a), a \wedge b, \Box \neg(a \wedge b) \vdash c \mid \Gamma_1} (com_{\bigcirc}^{\mathcal{K}}) \\
\frac{\bigcirc(b \mid a), a \wedge b, \Box \neg(a \wedge b) \vdash c \mid \Gamma_1}{a \wedge b, \Box \neg(a \wedge b) \vdash c \mid \Gamma_1} (com_{\bigcirc}^{\mathcal{K}}) \\
\frac{a \wedge b, \Box \neg(a \wedge b) \vdash c \mid \Gamma_1}{\bigcirc(b \mid a), \bigcirc(c \mid a), b, \boxed{c} \vdash \bigcirc(c \mid a \wedge b)} (\bigcirc R^{\mathcal{K}}) \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), b, \boxed{c} \vdash \bigcirc(c \mid a \wedge b)}{\bigcirc(b \mid a), \bigcirc(c \mid a), \boxed{b} \vdash \bigcirc(c \mid a \wedge b)} (\bigcirc L^{\mathcal{K}}) \\
\frac{\bigcirc(b \mid a), \bigcirc(c \mid a), \boxed{b} \vdash \bigcirc(c \mid a \wedge b)}{\bigcirc(b \mid a), \bigcirc(c \mid a) \vdash \bigcirc(c \mid a \wedge b)} (\bigcirc L^{\mathcal{K}})
\end{array}$$

Where Γ_1 - Γ_4 abbreviate the following four saturated components:

$$\begin{aligned}\Gamma_1 &= \bigcirc(b \mid a), \bigcirc(c \mid a), b, c \vdash \bigcirc(c \mid a \wedge b), \\ \Gamma_2 &= \bigcirc(b \mid a), \bigcirc(c \mid a), a \wedge b, a, b, \Box \neg(a \wedge b) \vdash c, \Box \neg a, \\ \Gamma_3 &= \bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, a, c, \neg(a \wedge b) \vdash \neg a, a \wedge b, b, \Box \neg a, \\ \Gamma_4 &= \bigcirc(b \mid a), \bigcirc(c \mid a), \blacktriangle, a, b, c, \Box \neg a \vdash \neg a.\end{aligned}$$

Taking the four components of the saturated hypersequent as worlds, the reconstructed valuation is

$$V(\Gamma_1) = \{b, c\}, \quad V(\Gamma_2) = \{a, b\}, \quad V(\Gamma_3) = \{a, c\}, \quad V(\Gamma_4) = \{a, b, c\}.$$

The strict part of the reconstructed preference relation is

$$\Gamma_1 \prec \Gamma_3, \quad \Gamma_1 \prec \Gamma_4, \quad \Gamma_2 \prec \Gamma_3, \quad \Gamma_3 \prec \Gamma_4,$$

with no other strict comparisons. In particular, $\Gamma_2 \prec \Gamma_3 \prec \Gamma_4$ while $\Gamma_2 \not\prec \Gamma_4$, illustrating the non-transitive structure permitted by the semantic characterization of **F**.

The unique best a -world is Γ_4 , so $\bigcirc(b \mid a)$ and $\bigcirc(c \mid a)$ both hold. The best $a \wedge b$ -worlds are Γ_2 and Γ_4 , and $\Gamma_2 \not\models c$, so $\bigcirc(c \mid a \wedge b)$ is false. Hence the reconstructed model falsifies the initial hypersequent.

4.9 Bridge: Embeddings of Åqvist's logics into normal bimodal logics

As in the case of **PCL**, the connection between Åqvist's logics and provability logics can be exploited not only proof-theoretically, but also semantically, yielding embeddings of Åqvist's systems into suitable bimodal logics.

For normal modal logics L_1 and L_2 defined by their frame properties, bimodal logic L_1+L_2 is defined over the language with two modalities $\Box_{L_1}$ and $\Box_{L_2}$:

$$\mathcal{L}_{L_1+L_2} = A ::= x \in \text{Atm} \mid \neg A \mid A \wedge A \mid A \vee A \mid A \rightarrow A \mid \Box_{L_1} A \mid \Box_{L_2} A.$$

L_1+L_2 can be defined semantically via relational models with two accessibility relations with corresponding frame properties. Note that we use finite-model characterization to make it easily applicable to provability logics too.

Definition 28. A L_1+L_2 -model is a tuple $M = \langle W, R_{L_1}, R_{L_2}, V \rangle$, where W is a finite set of worlds, R_{L_1} and R_{L_2} are binary relations on W , satisfying frame properties of L_1 and L_2 respectively, and V is valuation on W . Satisfaction of formulas from $\mathcal{L}_{L_1+L_2}$

w.r.t. L_1+L_2 -models is defined as follows:

$$\begin{array}{ll}
 M, w \models x & \text{iff } x \in V(w) \\
 M, w \models \neg A & \text{iff } M, w \not\models A \\
 M, w \models A \wedge B & \text{iff } M, w \models A \text{ and } M, w \models B \\
 M, w \models A \vee B & \text{iff } M, w \models A \text{ or } M, w \models B \\
 M, w \models A \rightarrow B & \text{iff } M, w \not\models A \text{ or } M, w \models B \\
 M, w \models \Box_{L_1} A & \text{iff } M, u \models A \text{ for any } u \in W \text{ s.t. } w R_{L_1} u \\
 M, w \models \Box_{L_2} A & \text{iff } M, u \models A \text{ for any } u \in W \text{ s.t. } w R_{L_2} u
 \end{array}$$

We can generalize the embedding of **E** into bimodal logic from [COP22], represent obligations using two interacting modalities: one governing the alethic accessibility relation and another governing the preference relation. This leads to the following family of translations $tr_{L_1+L_2}^{\mathbb{A}}$ from $\mathcal{L}_{\mathbb{A}q}$ to $\mathcal{L}_{L_1+L_2}$:

$$\begin{array}{ll}
 tr_{L_1+L_2}^{\mathbb{A}}(x) & = x \\
 tr_{L_1+L_2}^{\mathbb{A}}(\neg A) & = \neg tr_{L_1+L_2}^{\mathbb{A}}(A) \\
 tr_{L_1+L_2}^{\mathbb{A}}(A \wedge B) & = tr_{L_1+L_2}^{\mathbb{A}}(A) \wedge tr_{L_1+L_2}^{\mathbb{A}}(B) \\
 tr_{L_1+L_2}^{\mathbb{A}}(A \vee B) & = tr_{L_1+L_2}^{\mathbb{A}}(A) \vee tr_{L_1+L_2}^{\mathbb{A}}(B) \\
 tr_{L_1+L_2}^{\mathbb{A}}(A \rightarrow B) & = tr_{L_1+L_2}^{\mathbb{A}}(A) \rightarrow tr_{L_1+L_2}^{\mathbb{A}}(B) \\
 tr_{L_1+L_2}^{\mathbb{A}}(\Box A) & = \Box_{L_1} tr_{L_1+L_2}^{\mathbb{A}}(A) \\
 tr_{L_1+L_2}^{\mathbb{A}}(\Box(B|A)) & = \Box_{L_1}((tr_{L_1+L_2}^{\mathbb{A}}(A) \wedge \Box_{L_2} \neg tr_{L_1+L_2}^{\mathbb{A}}(A)) \rightarrow tr_{L_1+L_2}^{\mathbb{A}}(B))
 \end{array}$$

where L_1 captures the alethic modality and L_2 captures preference relation through the modality corresponding to $\boxtimes$.

As observed in [COP22], instantiating L_1 with **S5** and L_2 with **K**, yields an embedding of **E** into **S5+K**. Stronger Åqvist systems can be obtained by strengthening the second modal logic in accordance with the corresponding properties of the preference relation. Since **F_{CM}** is characterized by finite transitive preferential models, replacing **K** with **GL** immediately yields an embedding of **F_{CM}**.

Theorem 21. *A formula A is true in **F_{CM}** iff $tr_{\mathbf{S5+GL}}^{\mathbb{A}}(A)$ is true in **S5+GL**.*

Proof. From left to right, we need to verify that application of $tr_{\mathbf{S5+GL}}^{\mathbb{A}}(\cdot)$ to every rule of $\mathcal{H}_{\mathbf{F_{CM}}}$ provides a rule sound w.r.t. **S5+GL**. This verification coincides with the soundness proof in Th. 17. From right to left, we argue by contraposition and transform an arbitrary **F_{CM}**-countermodel $M = \langle W, \succeq, V \rangle$ into a **S5+GL**-countermodel $M' = \langle W, R_{\mathbf{S5}}, R_{\mathbf{GL}}, V \rangle$ with universal relation $R_{\mathbf{S5}}$ and $w R_{\mathbf{GL}} u$ iff $u \succ w$. It can be verified by trivial induction that $M, w \models A$ iff $M', w \models tr_{\mathbf{S5+GL}}^{\mathbb{A}}(A)$. $\square$

By Lem. 14, **G** is characterized by finite preferential models whose strict preference relation is transitive, asymmetric, and modular. Accordingly, the second modality must be interpreted in the normal modal logic determined by finite frames satisfying precisely

these properties. We denote this logic by $\mathbf{GL}_{\text{mod}}$. The logic $\mathbf{GL}_{\text{mod}}$ was already formally introduced in [BHM09] as a modal basis for conditional reasoning. Substituting L_2 in the above translation with $\mathbf{GL}_{\text{mod}}$ yields an embedding of $\mathbf{G}$.

Theorem 22. *Formula A is true in $\mathbf{G}$ iff $\text{tr}_{\mathbf{S5}+\mathbf{GL}_{\text{mod}}}^{\mathbf{A}}(A)$ is true in $\mathbf{S5}+\mathbf{GL}_{\text{mod}}$.*

Proof. Analogous to the proof of Th. 21, using the semantical characterization in Lem. 14. Modularity condition ensures soundness of $(\text{com}_{\Box})$ and corresponds to modularity of accessibility relation $R_{\mathbf{GL}_{\text{mod}}}$. $\square$

Finally, an analogous embedding can be obtained for $\mathbf{F}$ using our characterization with tail-transitivity from Sec. 4.8. In this case, we take as L_2 the normal modal logic $\mathbf{GL}_{\text{tail}}$ determined by finite frames whose accessibility relation is asymmetric and tail-transitive. This logic can be viewed as a weakening of $\mathbf{GL}$, obtained by replacing transitivity with tail-transitivity. To our knowledge, it does not coincide with any logic already studied in the literature. Nevertheless, we can use it to obtain the same embedding for $\mathbf{F}$.

Theorem 23. *Formula A is true in $\mathbf{F}$ iff $\text{tr}_{\mathbf{S5}+\mathbf{GL}_{\text{tail}}}^{\mathbf{A}}(A)$ is true in $\mathbf{S5}+\mathbf{GL}_{\text{tail}}$.*

Proof. Analogous to the proof of Th. 21, using the semantical characterization in Lem. 15. Tail-transitivity condition ensures soundness of $(\Box_{\mathbf{F}})$ and $(\Box_{\mathbf{F}}^{\mathbf{A}})$ and corresponds to tail-transitivity of accessibility relation $R_{\mathbf{GL}_{\text{tail}}}$. $\square$

Taken together, these embeddings reveal a close correspondence between the semantic hierarchy of Åqvist's logics and a hierarchy of normal modal logics governing the preference relation. In particular, the progression

$$\mathbf{E} \subseteq \mathbf{F} \subseteq \mathbf{F}_{\text{CM}} \subseteq \mathbf{G}$$

mirrors the progression

$$\mathbf{K} \subseteq \mathbf{GL}_{\text{tail}} \subseteq \mathbf{GL} \subseteq \mathbf{GL}_{\text{mod}}$$

providing an alternative modal perspective on the structure of the Åqvist hierarchy.

Norm-Based Deontic Logics

This chapter studies Input/Output (I/O) logic, one of the central representatives of the norm-based approach to deontic reasoning. Unlike preference-based semantics considered in the previous chapters, norm-based approaches do not assign truth conditions to norms themselves. Instead, I/O logic treat a normative theory as a transformation mechanism mapping factual contexts to their corresponding obligations. Over time, I/O logic evolved into a diverse family of different systems, applied to problems such as normative change and norm conflict resolution, and also used as the basis for causal and legal reasoning.

The chapter focuses on the twelve core I/O logics: the original systems $\mathbf{OUT}_1$ – $\mathbf{OUT}_4$, their causal counterparts $\mathbf{OUT}_1^\perp$ – $\mathbf{OUT}_4^\perp$, and the throughput variants $\mathbf{OUT}_1^+$ – $\mathbf{OUT}_4^+$. We begin by surveying the history and applications of I/O logics (Section 5.1) and the state of the art on their various characterizations (Section 5.2). We then introduce the twelve considered core I/O logics axiomatically via the inference principles they validate (Section 5.3). In Sec. 5.4 we develop sequent calculi for all these logics, starting with the causal versions, and then adapting them to the original and throughput versions. The development of these calculi is closely intertwined with a characterization of derivability, which we call *separation principles*, describing entailment in terms of partitions of normative/causal theories into two parts. In Sec. 5.5, we use this principles to establish a simple uniform reduction from every original I/O logic to its causal counterpart via a syntactic “bridging condition.” In Sec. 5.6, using contrapositive reading of the separation principles we obtain a new semantic characterization of all core logics in terms of a class of possible-world models, called *input/output models*. This semantics generalizes both the operational semantics of the original logics and the bimodel semantics of the causal logics. Using it, we construct simple embeddings of all core logics into classical and normal modal logics (Sec. 5.8) and reduce entailment in Åqvist’s logic $\mathbf{E}$ to entailment in I/O logics $\mathbf{OUT}_1$ and $\mathbf{OUT}_3$ (Sec. 5.9). The former embedding into classical logic serves as the basis for the automated prover `IOentailment` for original and causal I/O logics, developed by Marit Einböck, presented in Sec. 5.10. Finally, in Sec. 5.11 we apply

our calculi to the computational analysis of I/O logics and obtain a complete complexity classification for the full language and for important fragments, namely Horn conditionals and literal conditionals, which play a central role in causal applications. In addition, we show that Bochman’s notion of actual causes defined on the basis of $\mathbf{OUT}_3^\perp$ can be effectively computed in NP time using our calculi (Sec. 5.12).

5.1 Background

The preference-based tradition in deontic logic, explored in the previous chapter, relies on possible-world semantics to rank states of affairs according to their “ideality.” While powerful, this approach assumes that normative statements can be assigned truth values within a possible-world semantics, effectively treating them as descriptions of a preferred state. In contrast, the norm-based approach treats a normative system not as a set of true/false propositions, but as a formal mechanism for mapping factual contexts to their corresponding obligations. By decoupling the logic of the norms from the logic of the underlying propositions, this perspective avoids the complex philosophical and logical issues of assigning truth to imperatives—often referred to as Jørgensen’s Dilemma [J37]—and provide a more natural framework for defeasible normative reasoning. In norm-based systems, a normative code acts as a “black box” that can be updated or constrained without needing to re-rank an entire semantic model. This allows for a more efficient handling of exceptions and conflicts, making it a robust alternative for modeling real-world normative systems where requirements are frequently revised or overridden by new information.

The most prominent representative of this tradition is Input/Output (I/O) logic, introduced by Makinson and van der Torre [MvdT00]. Rather than treating norms as modal formulas, I/O logic represents them as ordered pairs (A, X) of propositional formulas, where A is the input condition and X is the mandated output obligation. Over the last two decades, I/O logic framework has evolved into a diverse family of different systems and has been used to provide a new perspective on a variety of important normative reasoning tasks. In particular, it was used as the basis for modeling different kinds of permission [MvdT03, OPvdT23], for modelling normative change in multi-agent systems [BPvdT09], for resolution of normative conflicts and dilemmas via integration with formal argumentation [AvBS24, vBS25], and for formalization of Kelsenian legal theories involving hierarchies of norms [CPS21]. Furthermore, the framework’s practical utility was demonstrated in modern Legal Tech, where it has been applied to the formal modeling and automated compliance verification of complex regulatory frameworks, such as EU’s General Data Protection Regulation (GDPR) [RBP⁺20].

Beyond normative reasoning, the Input/Output framework has found fundamental application in the logical formalization of causality. Bochman [Boc03] observed that the inference rules of I/O logic are well suited to characterize causal production relations, where the input and output of a pair represent cause and effect. By supplying his version of I/O logic with a fixed-point semantics, he established a natural logical framework [Boc04,

Boc21] with profound connections to various causal notions and formalisms. In particular, it provided an inferential characterization for both abductive reasoning [Boc07] and the causal calculus of McCain and Turner [MT97], served as the basis for a new well-behaved definition of actual causes [Boc18], and formalized core rationality postulates, including a Leibniz’s Principle of Sufficient Reason [Boc23].

As the preceding overview suggests, “Input/Output logic” is not a single system, but rather a broad family of logics defined by their specific structural properties. The original paper on I/O logic [MvdT00] introduced the “diamond” of four systems, **OUT**₁–**OUT**₄, which are distinguished by validation of principles of reasoning by cases (Or) and the reusability of outputs as inputs (CT). The companion paper [MvdT01] introduced the throughput variants **OUT**₁⁺–**OUT**₄⁺, extending the original diamond with the principle (Id) of passing inputs through to outputs; two of these — **OUT**₂⁺ and **OUT**₄⁺ — were shown to coincide with classical logic. The causal applications are based on yet another diamond **OUT**₁[⊥]–**OUT**₄[⊥] [Boc04], extending the original systems with the principle of explicit inconsistency handling. Many other variations have since emerged, e.g. systems for output aggregation [PvdT14] and systems with an integrated consistency-checking mechanism [PvdT18]. All these systems are monotonic, in the sense that the derivability of an I/O pair is preserved under extension of the normative (or causal) theory. Applications of I/O logics often require a non-monotonic layer built on top of these systems, and there are several different ways to do it depending on the application: constrained I/O logics [MvdT01] utilize an AGM-inspired approach by cutting the normative set at maxiconsistent subtheories, the causal application employs fixed-point semantics [Boc04], conflict resolution relies on the notion of stable extension of abstract argumentation frameworks [vBS25], and so on. To provide a clear formal treatment amidst this diversity, this chapter focuses on the twelve core systems (original, throughput, and causal ones) on the monotonic layer, in the hope that analytic characterizations of this fundamental core will serve as a solid foundation for studying other variations and non-monotonic extensions.

5.2 State of the art

There is a rich body of work studying Input/Output logics from a variety of semantic, proof-theoretic, and computational perspectives. Many different characterizations and results have been obtained; however, they often apply only to particular fragments or variants of the framework and do not extend uniformly to the whole family, leaving a number of natural questions open. The original paper on I/O logics [MvdT00] and its companion paper on constrained versions [MvdT01], introduced original and throughput I/O logics both axiomatically and via an operational semantics based on mappings between sets of propositional formulas. The causal variants were obtained by extending these axiomatizations with the additional principle (Bot), but their semantics is formulated in a rather different way, using as models pairs of deductively closed sets of formulas (called *bimodels*) [Boc04]. Despite the close similarity of the axiomatizations, the precise relation between operational semantics and bimodel semantics remained unclear and was explicitly

posed as an open problem in the theory of I/O logics [vdTP13]. The original paper also showed that **OUT₂** and **OUT₄** admit natural embeddings into classical logic and the normal modal logics **K** and **T**, respectively. Similar embeddings were later obtained for their causal counterparts [Boc04], while the throughput counterparts collapse to classical logic [MvdT01]. However, the methodology used for these embeddings did not extend to the remaining systems in the family, leading even to the conjecture that embeddings for **OUT₁** and **OUT₃** are not possible “in a straightforward way” [MvdT00]. Other semantic approaches include fixed-point semantics for **OUT₃** [Sun14], embeddings into Formal Concept Analysis framework for **OUT₂** and **OUT₄** [Sto15], characterization in terms subordination algebras [DFM⁺22] and neighbourhood models [Hop25].

Several proof-theoretic approaches to I/O logics have also been proposed. Proof calculi for original and throughput I/O logics (as well as their constrained variants) were developed in the highly expressive framework of adaptive logics [SBP16]. Sequent calculi were later proposed for several I/O logics and connected with argumentation frameworks to model different mechanisms for resolving conflicts between obligations [vBS22, AvBRS24]. These calculi, however, are not analytic and do not naturally support proof search. Analytic sequent-style calculi, based on “sequents with blocks”, were developed for a number of I/O logics, including **OUT₁** and **OUT₃** [Lel21]. The proposed methodology does not extend, however, to **OUT₂** and **OUT₄**.

On the computational side, entailment in most original and throughput I/O logics was shown to be co-NP-complete [SR17]. A notable exceptions are **OUT₃** and **OUT₃⁺**, where only a co-NP lower bound and a P^{NP} upper bound were established. Several methods for automated deduction have also been proposed and implemented. One approach is based on proof search in the above-mentioned sequent calculi for **OUT₁** and **OUT₃**. Another approach [BFMP19] implements embeddings of **OUT₂** and **OUT₄** into normal modal logics within Isabelle/HOL framework. More recently, a SAT-based approach covering all original and throughput I/O logics (and their constrained variants) was proposed [Ste25], where the most general output entailed by a given input with respect to a normative base is computed. However, these approaches were not aimed at efficient entailment checking, and little complexity analysis of them has been carried out: the first approach was remarked to solve entailment in the class Π_3^P , well above the known theoretical bounds, while the last approach performs polynomially many SAT queries for **OUT₁** and **OUT₃**, but for **OUT₂** and **OUT₄** the computed most general output may have exponential size in the worst case.

5.3 Preliminaries: Axioms of the core I/O logics

In the I/O logic framework, conditionals are expressed as pairs of propositional Boolean formulas. The core I/O logics (as well as various other related systems) can be defined axiomatically via the derivation rules, governing how new input-output pairs can be derived from a given set of these pairs. For the core I/O logics these rules, that we will refer to as *I/O principles*, are the given below ($Cn(\mathcal{A})$ denotes the set of classical

consequences of a set $\mathcal{A}$ of formulas).

- (**Top**) $(\top, \top)$ is derivable from no premises
- (**Bot**) $(\perp, \perp)$ is derivable from no premises
- (**Id**) (A, A) is derivable from no premises
- (**WO**) (A, X) derives (A, Y) whenever $Y \in Cn(\{X\})$
- (**SI**) (A, X) derives (B, X) whenever $B \in Cn(\{A\})$
- (**And**) (A, X_1) and (A, X_2) derive $(A, X_1 \wedge X_2)$
- (**Or**) (A_1, X) and (A_2, X) derive $(A_1 \vee A_2, X)$
- (**CT**) (A, X) and $(A \wedge X, Y)$ derive (A, Y)

Different I/O logics are given by different subsets R of these principles, shown in Tab. 5.1. The basic system, called *simple-minded output* **OUT**₁, consists of the rules $\{(\text{Top}), (\text{WO}), (\text{SI}), (\text{And})\}$. Its extension with (Or) (for reasoning by cases) leads to *basic output* logic **OUT**₂, with (CT) (for reusability of outputs as inputs in derivations) to *simple-minded reusable output* logic **OUT**₃, and with both (Or) and (CT) to *basic reusable output* logic **OUT**₄. We will refer to these logics as the original I/O logics. Their causal counterpart [Boc04], that we denote by **OUT** _{i} [⊥] for $i = 1, \dots, 4$, extend the corresponding **OUT** _{i} with (Bot). The throughput I/O logics **OUT**₁⁺-**OUT**₄⁺ [MvdT01] are obtained by extending these systems with (Id), which enforces that every accepted input is also accepted as an output.

Definition 29. *Given a set of pairs G and a set R of principles, a derivation in an I/O logic of a pair (B, Y) from G is a tree with (B, Y) at the root, each non-leaf node derivable from its immediate parents by one of the principles in R , and each leaf node is an element of G or a pair derived by one of the principles in R with no premises.*

We indicate by $G \vdash_{\mathcal{L}} (B, Y)$ that the pair (B, Y) is derivable from the set of pairs in G in the I/O logic $\mathcal{L}$. We will refer to (B, Y) as the *goal pair*, to the formulas B and Y as the *goal input* and *goal output* respectively, and to the pairs in G as *deriving pairs*.

A derivation in I/O logic is a kind of natural deduction proof, which operates on pairs rather than on single formulas. However this proof theory is not analytic and not very useful for determining whether $G \vdash_{\mathcal{L}} (B, Y)$ holds, nor for establishing metalogical properties such as complexity bounds. The main reason is that these derivations do not admit well-behaved normal forms, and are in general difficult to construct.

Logic	(Top)	(Bot)	(Id)	(WO)	(SI)	(And)	(Or)	(CT)
OUT₁	✓			✓	✓	✓		
OUT₂	✓			✓	✓	✓	✓	
OUT₃	✓			✓	✓	✓		✓
OUT₄	✓			✓	✓	✓	✓	✓
OUT₁[⊥]	✓	✓		✓	✓	✓		
OUT₂[⊥]	✓	✓		✓	✓	✓	✓	
OUT₃[⊥]	✓	✓		✓	✓	✓		✓
OUT₄[⊥]	✓	✓		✓	✓	✓	✓	✓
OUT₁⁺	✓	✓	✓	✓	✓	✓		
OUT₂⁺	✓	✓	✓	✓	✓	✓	✓	
OUT₃⁺	✓	✓	✓	✓	✓	✓		✓
OUT₄⁺	✓	✓	✓	✓	✓	✓	✓	✓

Table 5.1: Core I/O logics defined by I/O principles

5.4 Sequent calculi and separation principles for I/O logics

In this section, we introduce uniform and modular sequent calculi for all core I/O logics, with proof search proceeding stepwise by eliminating all deriving pairs one by one. We first define the calculi for the causal I/O logics, followed by the original and throughput versions. In the process of proving completeness of the calculi, we establish an alternative characterization of derivability in these calculi. It is based on splitting the set of deriving pairs in $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ into two parts and reducing the derivability in our calculi to the derivability in classical logic for each partition. This characterization, which we will refer to as *separation principles*, is illustrated for all considered logics in Tab. 5.2. It will play a key role not only in proving the admissibility of the structural rules, which we will use to establish the completeness of the calculi, but also in development of the alternative semantics later on.

Our calculi manipulate two types of sequents: *I/O sequents* of the form

$$(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y),$$

capturing entailment of a pair in a given I/O logic, as well as standard **LK** sequents

$$A_1, \dots, A_n \vdash B_1, \dots, B_m$$

where $A_1, \dots, A_n, B_1, \dots, B_m$ are Boolean formulas.

All calculi are defined in a modular way, by extending the sequent calculus **LK** for classical logic (Sec. 2.2) with rules that manipulate I/O pairs. These additional rules are divided into three groups: the *transition rules* (*IN*), (*OUT*) and (*THROU*), which reduce the derivation of the goal pair to an **LK** derivation; the *elimination rules*, which remove one of the deriving pair while modifying the goal pair; and the *structural rules* of weakening, contraction and cut, applied to pairs. As we will show, structural rules are admissible in all calculi except for those for **OUT₂** and **OUT₄**. In these two cases, contraction, which

Transition rules:

$$\frac{B \vdash}{G \vdash (B, Y)} \text{ (IN)} \quad \frac{\vdash Y}{G \vdash (B, Y)} \text{ (OUT)} \quad \frac{B \vdash Y}{G \vdash (B, Y)} \text{ (THROU)}$$

Pair-elimination rules:

$$\begin{array}{c} \frac{B \vdash A \quad G \vdash (B, Y \vee \neg X)}{(A, X), G \vdash (B, Y)} (E_1) \\ \frac{B \vdash A \quad G \vdash (B \wedge X, Y \vee \neg X)}{(A, X), G \vdash (B, Y)} (E_3) \end{array} \quad \begin{array}{c} \frac{G \vdash (B \wedge \neg A, Y) \quad G \vdash (B, Y \vee \neg X)}{(A, X), G \vdash (B, Y)} (E_2) \\ \frac{G \vdash (B \wedge \neg A, Y) \quad G \vdash (B \wedge X, Y \vee \neg X)}{(A, X), G \vdash (B, Y)} (E_4) \end{array}$$

Structural rules:

$$\begin{array}{c} \frac{G \vdash (B, Y)}{(A, X), G \vdash (B, Y)} \text{ (IO-Wk)} \quad \frac{(A, X), (A, X), G \vdash (B, Y)}{(A, X), G \vdash (B, Y)} \text{ (IO-Ctr)} \\ \frac{G \vdash (C, Z) \quad (C, Z), G' \vdash (B, Y)}{G, G' \vdash (B, Y)} \text{ (IO-Cut)} \end{array}$$

$$\begin{array}{l} SC_{\text{OUT}_1} = \{(\text{OUT}), (E_1)\} \\ SC_{\text{OUT}_1^\perp} = \{(\text{IN}), (\text{OUT}), (E_1)\} \\ SC_{\text{OUT}_1^+} = \{(\text{THROU}), (E_1)\} \\ SC_{\text{OUT}_2} = \{(\text{OUT}), (E_1), (E_2), (\text{IO-Ctr})\} \\ SC_{\text{OUT}_2^\perp} = \{(\text{IN}), (\text{OUT}), (E_2)\} \\ SC_{\text{OUT}_2^+} = \{(\text{THROU}), (E_2)\} \\ SC_{\text{OUT}_3} = \{(\text{OUT}), (E_3)\} \\ SC_{\text{OUT}_3^\perp} = \{(\text{IN}), (\text{OUT}), (E_3)\} \\ SC_{\text{OUT}_3^+} = \{(\text{THROU}), (E_3)\} \\ SC_{\text{OUT}_4} = \{(\text{OUT}), (E_3), (E_4), (\text{IO-Ctr})\} \\ SC_{\text{OUT}_4^\perp} = \{(\text{IN}), (\text{OUT}), (E_4)\} \\ SC_{\text{OUT}_4^+} = \{(\text{THROU}), (E_4)\} \end{array}$$

Figure 5.1: Sequent calculi for the core I/O logics

Logic	Input sequent(s)	Ex-falso sequent	Output sequent
OUT₁	$\{B \vdash A_i\}_{i \in I}$	-	$\{X_j\}_{j \in J} \vdash Y$
OUT₁[⊥]	$\{B \vdash A_i\}_{i \in I}$	$B \vdash \perp$	$\{X_j\}_{j \in J} \vdash Y$
OUT₁⁺	$\{B \vdash A_i\}_{i \in I}$	-	$B, \{X_j\}_{j \in J} \vdash Y$
OUT₂	$B \vdash \{A_i\}_{i \in I} \quad \text{if } I \neq \emptyset$	-	$\{X_j\}_{j \in J} \vdash Y$
OUT₂[⊥]	$B \vdash \{A_i\}_{i \in I}$	-	$\{X_j\}_{j \in J} \vdash Y$
OUT₂⁺	-	-	$B, \{X_j\}_{j \in J} \vdash Y, \{A_i\}_{i \in I}$
OUT₃	$\{B, \{X_j\}_{j \in J} \vdash A_i\}_{i \in I}$	-	$\{X_j\}_{j \in J} \vdash Y$
OUT₃[⊥]	$\{B, \{X_j\}_{j \in J} \vdash A_i\}_{i \in I}$	$B, \{X_j\}_{j \in J} \vdash \perp$	$\{X_j\}_{j \in J} \vdash Y$
OUT₃⁺	$\{B, \{X_j\}_{j \in J} \vdash A_i\}_{i \in I}$	-	$B, \{X_j\}_{j \in J} \vdash Y$
OUT₄	$B, \{X_j\}_{j \in J} \vdash \{A_i\}_{i \in I} \quad \text{if } I \neq \emptyset$	-	$\{X_j\}_{j \in J} \vdash Y$
OUT₄[⊥]	$B, \{X_j\}_{j \in J} \vdash \{A_i\}_{i \in I}$	-	$\{X_j\}_{j \in J} \vdash Y$
OUT₄⁺	-	-	$B, \{X_j\}_{j \in J} \vdash Y, \{A_i\}_{i \in I}$

Table 5.2: Separation principles for the core I/O logics. General form: $(A_1, X_n), \dots, (A_n, X_n) \vdash_{\mathcal{L}} (B, Y)$ iff for all partitions $(I, J) \in \mathcal{P}(\{1, \dots, n\})$ (where $\mathcal{P}(X) = \{(I, J) \mid I \cup J = X, I \cap J = \emptyset\}$) either the *output sequent*, the *ex-falso sequent* (if any) or one of the *input sequents* is derivable in **LK**.

duplicates a deriving pair, is essential. The complete set of rules and the structure of the calculi are presented in Figure 5.1.

A derivation of an I/O sequent consists of two phases. Looking at it bottom up, in each branch we first encounter rules dealing with pairs (pair elimination rules, and for **OUT₂** and **OUT₄** also contraction), then a transition rule, followed by **LK** rules. The calculi, in a sense, uphold the ideological principles guiding I/O logics: pairs (e.g., conditional norms) are treated separately from the boolean statements. Henceforth, when presenting derivations in our calculi, we will omit the **LK** sub-derivations.

5.4.1 Sequent Calculi for Causal I/O logics

All calculi for the causal I/O logics **OUT₁[⊥]**-**OUT₄[⊥]** follow the same pattern: they consist of a core calculus, which is **LK** augmented with the transition rules (IN) and (OUT) from Fig. 5.1, and one *elimination rule*, different for each of the logic. It is easy to see that using (IN) and (OUT) we can derive (TOP) and (BOT), respectively; conversely we can derive these rules via I/O principles of **OUT₁[⊥]** (and hence of all causal I/O logics).

Lemma 16. (IN) and (OUT) are derivable in **OUT₁[⊥]**.

Proof. For (IN): Assume the premise $B \vdash$ is derivable in **LK**, then $\perp \in Cn(\{B\})$. From $(\perp, \perp)$ and $\perp \in Cn(\{B\})$ we get $(B, \perp)$ by (SI); the required pair (B, Y) follows by (WO). For (OUT): Assume the premise $\vdash Y$ is derivable in **LK**, then $Y \in Cn(\{\top\})$.

From $(\top, \top)$ and $Y \in Cn(\{\top\})$ by (WO) we get $(\top, Y)$, from which (B, Y) follows by (SI) . $\square$

We begin by presenting the calculi for the causal I/O logics which include the rule (OR) .

Sequent calculus for $\mathbf{OUT}_2^\perp$

The calculus $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$ for the causal basic output logic $\mathbf{OUT}_2^\perp$ is obtained by adding to the core calculus the pair elimination rule (E_2) depicted in Fig. 5.1.

Remark 5. *The rule (E_2) takes inspiration from the following embedding of $\mathbf{OUT}_2^\perp$ into the modal logic $\mathbf{K}$ from [Boc04]: $(A_1, X_1), \dots, (A_n, X_n) \vdash_{\mathbf{OUT}_2^\perp} (B, Y)$ iff sequent $(*) (A_1 \rightarrow \Box X_1), \dots, (A_n \rightarrow \Box X_n), B \vdash \Box Y$ is valid in logic $\mathbf{K}$. To provide the rule's intuition we make use of the sequent calculus $\mathcal{SC}_{\mathbf{K}}$ for $\mathbf{K}$ presented in Sec. 2.2 (Fig. 2.3). $\mathcal{SC}_{\mathbf{K}}$ extends $\mathbf{LK}$ with the rule $(\Box_{\mathbf{K}}R)$ for eliminating $\Box$ -modalities. To find a derivation for the sequent $(*)$ in $\mathcal{SC}_{\mathbf{K}}$ we start bottom-up by applying the $\mathbf{LK}$ rule for $(\rightarrow L)$ to one of the implications $(A_i \rightarrow \Box X_i)$ on the left. This leads to two premises: (a) $G', B \vdash A_i, \Box Y$ and (b) $\Box X_i, G', B \vdash \Box Y$ (where G' is the set of all implications on the left-hand side apart from $(A_i \rightarrow \Box X_i)$). Now (a) is equivalent to (i.e. derivable in $\mathcal{SC}_{\mathbf{K}}$ if and only if so is) the sequent $G', (B \wedge \neg A_i) \vdash \Box Y$, that using the embedding again leads to the first premise of (E_2) ; (b) is equivalent (for this form of G') to $G', B \vdash \Box(Y \vee \neg X_i)$, and this leads to the second premise of (E_2) .*

Theorem 24 (Soundness of $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$). *If $G \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$, then $G \vdash_{\mathbf{OUT}_2^\perp} (B, Y)$.*

Proof. We can represent every sequent rule of $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$ as a sequence of applications of rules in $\mathbf{OUT}_2^\perp$. A simulation of (E_2) is as follows:

$$\begin{array}{c}
 \frac{(A, X)}{(A, Y \vee X)} \text{ (WO)} \\
 \frac{(A, Y \vee X)}{(B \wedge A, Y \vee X)} \text{ (SI)} \quad \frac{(B, Y \vee \neg X)}{(B \wedge A, Y \vee \neg X)} \text{ (SI)} \\
 \frac{}{(B \wedge A, (Y \vee X) \wedge (Y \vee \neg X))} \text{ (AND)} \\
 \frac{}{(B \wedge A, Y \vee (X \wedge \neg X))} \text{ (WO)} \\
 \frac{(B \wedge A, Y)}{(B \wedge A, Y)} \text{ (WO)} \quad \frac{(B \wedge \neg A, Y)}{(B \wedge \neg A, Y)} \text{ (OR)} \\
 \frac{}{((B \wedge A) \vee (B \wedge \neg A), Y)} \text{ (SI)} \\
 \frac{}{(B \wedge (A \vee \neg A), Y)} \text{ (SI)} \\
 \frac{}{(B, Y)} \text{ (SI)}
 \end{array}$$

The claim follows by the soundness of (IN) and (OUT) (Lem. 16). $\square$

We prove below the completeness of the calculus $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$ for $\mathbf{OUT}_2^\perp$. This relies on the admissibility of the structural rules for I/O sequents $((IO\text{-}Wk), (IO\text{-}Ctr), (IO\text{-}Cut))$

in Fig. 5.1. Recall that a rule is *admissible* if its addition does not change the set of sequents that can be derived.

We start by describing a useful characterization of derivability of an I/O sequent $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ in $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$ via derivability of two **LK**-sequents:

- an *input sequent* that contains only the input formulas, and
- an *output sequent* that contains only the output formulas

for every possible way of arranging how the inputs and outputs interact. To establish the admissibility of the structural rules, we make essential use of this characterization. $\mathcal{P}(X)$ will denote the set of all *partitions* of the set X , i.e., $\mathcal{P}(X) = \{(I, J) \mid I \cup J = X, I \cap J = \emptyset\}$.

Lemma 17 (Separation Principle for $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$). *$(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$ iff for all partitions $(I, J) \in \mathcal{P}(\{1, \dots, n\})$, either $B \vdash \{A_i\}_{i \in I}$ (input sequent) or $\{X_j\}_{j \in J} \vdash Y$ (output sequent) is derivable in **LK**.*

Proof. We argue by contraposition in both directions.

(*Left-to-right*). Suppose there exists a derivation of $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ in $\mathcal{SC}_{\mathbf{OUT}_2}$, and suppose towards contradiction that some partition (I, J) makes both associated **LK**-sequents underivable. Traverse the given derivation bottom-up as follows: For each application of the rule (E_2) to a pair (A_i, X_i) we move to the left premise if $i \in I$, and to the right premise if $j \in J$. Along this path the goal pair of the I/O-sequent accumulates the conjuncts $\wedge \neg A_i$ in the input if $i \in I$, and the disjunct $\vee \neg X_i$ in the output, if $i \in J$. Eventually, an application of a transition rule (either (IN) or (OUT)) is reached. Its goal pair must be of the form $(B \wedge \bigwedge_{i \in I'} \neg A_i, Y \wedge \bigvee_{j \in J'} \neg X_j)$, for some $I' \subseteq I$ and $J' \subseteq J$.

If the rule is (IN) , then the sequent $B \wedge \bigwedge_{i \in I'} \neg A_i \vdash$ is derivable in **LK**, and if the rule is (OUT) , then the sequent $\vdash Y \wedge \bigvee_{j \in J'} \neg X_j$ is derivable in **LK**. The derivability of the former sequent is equivalent (in **LK**) to the derivability of $B \vdash \{A_i\}_{i \in I'}$ and the derivability of the latter is equivalent to the derivability of $\{X_j\}_{j \in J'} \vdash Y$. By admissibility of weakening in **LK**, the former yields derivability of the input sequent for (I, J) and the latter that of the output sequent, a contradiction.

(*Right-to-left*). Suppose $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is not derivable in $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$. We construct a partition $(I, J) \in \mathcal{P}(\{1, \dots, n\})$ that makes both associated **LK**-sequents underivable. To achieve this, we iteratively transform the I/O sequent $G \vdash (B, Y)$ into another non-derivable sequent with less deriving pairs. Namely, at each iteration $i \in \{1, \dots, n\}$ we consider an application of the rule (E_2) , eliminating the pair (A_i, X_i) on the left. Since the conclusion is underivable, at least one premise is underivable; replace the current sequent with such a premise. After n such iterations, we obtain a non-derivable I/O sequent with no pairs on the left. Let I denote the set of indices for

which we took the left premise when eliminating the pair (A_i, X_i) , and J denote the set of indices for which we took the right one. The final non-derivable sequent we obtain is $\vdash (B \wedge \bigwedge_{i \in I} \neg A_i, Y \wedge \bigvee_{j \in J} \neg X_j)$. Neither the **LK** sequent $B \wedge \bigwedge_{i \in I} \neg A_i \vdash$ obtained by applying the rule (IN) , nor $\vdash Y \wedge \bigvee_{j \in J} \neg X_j$ obtained by the rule (OUT) , is derivable in **LK**; these two sequents are equivalent to the input and output sequents for the partition (I, J) . $\square$

The above separation principle can also be applied as a criterion of non-derivability: $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is *not* derivable in $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$ iff for some partition (I, J) neither $B \vdash \{A_i\}_{i \in I}$ nor $\{X_j\}_{j \in J} \vdash Y$ is derivable in **LK**.

Lemma 18. *The structural rules $(IO-Wk)$, $(IO-Ctr)$ and $(IO-Cut)$ in Fig. 5.1 are admissible in $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$.*

Proof. Using Lem. 17 we can reduce the admissibility of these structural rules to the admissibility of weakening, contraction, and cut in **LK**. For $(IO-Wk)$ and $(IO-Ctr)$ this is straightforward. We show below the admissibility of $(IO-Cut)$ in $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$.

Consider an instance of $(IO-Cut)$ with premises $G \vdash (A_0, X_0)$ and $(A_0, X_0), G' \vdash (B, Y)$, both derivable in $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$. We show that the conclusion $G, G' \vdash (B, Y)$ is derivable as well. Let $G = \{(D_1, W_1) \dots (D_m, W_m)\}$ and $G' = \{(A_1, X_1), \dots (A_n, X_n)\}$. By Lem. 17 the conclusion is derivable if for any $(I', J') \in \mathcal{P}(\{1, \dots, m\})$ and $(I'', J'') \in \mathcal{P}(\{1, \dots, n\})$ either the input sequent (InS) $B \vdash \{A_i\}_{i \in I''}, \{D_i\}_{i \in I'}$ or the output sequent (OutS) $\{X_j\}_{j \in J''}, \{W_j\}_{j \in J'} \vdash Y$ is derivable in **LK**. We proceed in three steps, applying Lem. 17 to both premises of cut for different partitions.

Step 1. We apply Lem. 17 to the premise $(A_0, X_0), (A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ and the partition $(I'' \cup \{0\}, J'') \in \mathcal{P}(\{0, 1, \dots, n\})$. Two cases can arise. The first is when the output sequent $\{X_j\}_{j \in J''} \vdash Y$ is derivable in **LK**, in which case (OutS) is also derivable in **LK** due to admissibility of weakening, and the claim holds. The remaining case is that the following input sequent, which we will use further, is derivable in **LK**

$$(1) \quad B \vdash A_0, \{A_i\}_{i \in I''}.$$

Step 2. We apply Lem. 17 to the premise $(D_1, W_1), \dots, (D_m, W_m) \vdash (A_0, X_0)$ and the partition $(I', J') \in \mathcal{P}(\{1, \dots, m\})$. Two cases can arise. The first is when the input sequent $A_0 \vdash \{D_i\}_{i \in I'}$ is derivable in **LK**; an application of cut with the sequent (1) in **LK** gives the sequent (InS), and the claim follows by the admissibility of the cut rule in **LK**. The remaining case is that the following output sequent, which we will use further, is derivable in **LK**

$$(2) \quad \{W_j\}_{j \in J'} \vdash X_0.$$

Step 3. We apply Lem. 17 to the premise $(A_0, X_0), (A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ and partition $(I'', J'' \cup \{0\}) \in \mathcal{P}(\{0, 1, \dots, n\})$. Two cases can arise. The first is when the input sequent $B \vdash \{A_i\}_{i \in I''}$ is derivable in **LK**, and the derivability of (InS) in **LK**

follows by the admissibility of weakening. The second case is that the output sequent $X_0, \{X_j\}_{j \in J''} \vdash Y$ is derivable in **LK**; the derivability of (OutS) follows by a cut in **LK** with (2). $\square$

Using the admissibility of the structural rules, we can now establish completeness of $SC_{\mathbf{OUT}_2^\perp}$

Theorem 25 (Completeness of $SC_{\mathbf{OUT}_2^\perp}$). *If $G \vdash_{\mathbf{OUT}_2^\perp} (B, Y)$ then $G \vdash (B, Y)$ is derivable in $SC_{\mathbf{OUT}_2^\perp}$.*

Proof. The proof proceeds by induction on the length of the derivation of (B, Y) in $\mathbf{OUT}_2^\perp$. The base case $((A, X) \in G)$ is handled as follows:

$$\frac{\frac{A \wedge \neg A \vdash}{G' \vdash (A \wedge \neg A, X)} (\text{IN}) \quad \frac{\vdash X \vee \neg X}{G' \vdash (A, X \vee \neg X)} (\text{OUT})}{(A, X), G' \vdash (A, X)} (E_2)$$

The simulation of (SI) is as follows (using the fact that $A \in Cn(\{B\})$ implies derivability the sequent $B \wedge \neg A \vdash$ in **LK**):

$$\frac{\begin{array}{c} \text{by I.H.} \\ \vdots \\ G \vdash (A, Y) \end{array} \quad \frac{\frac{B \wedge \neg A \vdash}{\vdash (B \wedge \neg A, Y)} (\text{IN}) \quad \frac{\vdash Y \vee \neg Y}{\vdash (B, Y \vee \neg Y)} (\text{OUT})}{(A, Y) \vdash (B, Y)} (E_2)}{G \vdash (B, Y)} (\text{IO-Cut})$$

The simulation of (WO) is as follows (using the fact that $Y \in Cn(\{X\})$ implies derivability of the sequent $\vdash Y \vee \neg X$ in **LK**):

$$\frac{\begin{array}{c} \text{by I.H.} \\ \vdots \\ G \vdash (B, X) \end{array} \quad \frac{\frac{B \wedge \neg B \vdash}{\vdash (B \wedge \neg B, Y)} (\text{IN}) \quad \frac{\vdash Y \vee \neg X}{\vdash (B, Y \vee \neg X)} (\text{OUT})}{(B, X) \vdash (B, Y)} (E_2)}{G \vdash (B, Y)} (\text{IO-Cut})$$

For (AND), we first derive sequent $(B, X_1), (B, X_2) \vdash (B, X_1 \wedge X_2)$ as follows:

$$\frac{\frac{B \wedge \neg B \vdash}{(B, X_2) \vdash (B \wedge \neg B, X_1 \wedge X_2)} (\text{IN}) \quad \frac{\frac{B \wedge \neg B \vdash}{\vdash (B \wedge \neg B, (X_1 \wedge X_2) \vee \neg X_1)} (\text{IN}) \quad \frac{\vdash (X_1 \wedge X_2) \vee \neg X_1 \vee \neg X_2}{\vdash (B, (X_1 \wedge X_2) \vee \neg X_1 \vee \neg X_2)} (\text{OUT})}{(B, X_2) \vdash (B, (X_1 \wedge X_2) \vee \neg X_1)} (E_2)}{(B, X_1), (B, X_2) \vdash (B, X_1 \wedge X_2)} (E_2)$$

Now we apply (IO-Cut) twice followed by possibly many applications of contraction to the resulting derivation as follows:

$$\frac{\begin{array}{c} \text{by I.H.} \\ \vdots \\ G \vdash (B, X_1) \end{array} \quad \frac{\frac{\vdots}{G \vdash (B, X_2)} \quad (B, X_1), (B, X_2) \vdash (B, X_1 \wedge X_2)}{G, (B, X_1) \vdash (B, X_1 \wedge X_2)} (\text{IO-Cut})}{G, G \vdash (B, X_1 \wedge X_2)} (\text{IO-Cut})} \frac{}{G \vdash (B, X_1 \wedge X_2)} (\text{IO-Ctr}) \times n$$

For (OR) we follow the same pattern as for (AND), with the following derivation for sequent $(A_1, Y), (A_2, Y) \vdash (A_1 \vee A_2, Y)$:

$$\frac{\frac{(A_1 \vee A_2) \wedge \neg A_1 \wedge \neg A_2 \vdash}{\vdash ((A_1 \vee A_2) \wedge \neg A_1 \wedge \neg A_2, Y)} \text{(IN)} \quad \frac{\vdash Y \vee \neg Y}{\vdash ((A_1 \vee A_2) \wedge \neg A_1, Y \vee \neg Y)} \text{(OUT)} \quad \frac{\vdash Y \vee \neg Y}{(A_2, Y) \vdash (A_1 \vee A_2, Y \vee \neg Y)} \text{(OUT)} \quad \frac{(A_2, Y) \vdash ((A_1 \vee A_2) \wedge \neg A_1, Y)}{(A_1, Y), (A_2, Y) \vdash (A_1 \vee A_2, Y)} \text{(E}_2\text{)}$$

The derivations of (BOT) and (TOP) are trivial. The claim follows by admissibility of (IO-Ctr) and (IO-Cut) in $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$ (Lem. 18). $\square$

Sequent Calculus for $\mathbf{OUT}_4^\perp$.

The calculus $\mathcal{SC}_{\mathbf{OUT}_4^\perp}$ for $\mathbf{OUT}_4^\perp$ extends the core calculus (consisting of **LK** with the rules (IN) and (OUT)) with the pair elimination rule (E_4) in Figure 5.1.

Remark 6. The rule (E_4) follows the same intuition as the rule for $\mathbf{OUT}_2^\perp$, but now the embedding is carried out in the modal logic **T** rather than in **K**. Under the embedding of [Boc04], the sequent $(A_1, X_1), \dots, (A_n, X_n) \vdash_{\mathbf{OUT}_4^\perp} (B, Y)$ corresponds to the validity of sequent $(*) (A_1 \rightarrow \Box X_1), \dots, (A_n \rightarrow \Box X_n), B \vdash \Box Y$ in **T**¹. Compared with $\mathcal{SC}_{\mathbf{K}}$, the calculus for $\mathcal{SC}_{\mathbf{T}}$ adds a rule ($\Box_{\mathbf{T}}L$) allowing one to drop a $\Box$ in the antecedent (Fig. 2.3). Applying the **LK** rule ($\rightarrow L$) to a formula $(A_i \rightarrow \Box X_i)$ in $(*)$ yields the two premises: (a) $G', B \vdash A_i, \Box Y$, and (b) $\Box X_i, G', B \vdash \Box Y$, with G' denoting the remaining implications. Premise (a) gives the first premise of (E_4). Premise (b) is equivalent in $\mathcal{SC}_{\mathbf{T}}$ to $X_i, \Box X_i, G', B \vdash \Box Y$, (by application of **LK**-contraction and ($\Box_{\mathbf{T}}L$) in one direction, and **LK**-weakening in the other), which in turn is equivalent in **T** (for formulas of this form in G') to $G', B \wedge X_i \vdash \Box(Y \vee \neg X_i)$, yielding the second premise of (E_4).

The shape of the rule (E_4) requires to amend the statement of the separation principle.

Lemma 19 (Separation Principle for $\mathcal{SC}_{\mathbf{OUT}_4^\perp}$). $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\mathbf{OUT}_4^\perp}$ iff for all $(I, J) \in \mathcal{P}(\{1, \dots, n\})$, either $B, \{X_j\}_{j \in J} \vdash \{A_i\}_{i \in I}$ (input sequent) or $\{X_j\}_{j \in J} \vdash Y$ (output sequent) is derivable in **LK**.

Proof. The argument proceeds as the one in the proof of Lemma 17. The only difference is what happens when we eliminate a pair (A_i, X_i) with the rule (E_4). In this case, moving to the right premise requires us to also extend the input component of the goal pair by conjoining it with $\wedge X_i$. As a consequence, in the *right-to-left* direction, the final goal pair will be $(B \wedge \bigwedge_{i \in I'} \neg A_i \wedge \bigwedge_{j \in J'} X_j, Y \wedge \bigvee_{j \in J'} \neg X_j)$. The same pattern holds for the *left-to-right* direction, except that the sets I' and J' are replaced by I and J , resulting in the addition of $\{X_j\}_{j \in J}$ on the left in the input sequent. $\square$

¹Actually, [Boc04] uses the relaxed property of quasi-reflexivity rather than reflexivity to establish a more general correspondence between Kripke models and set of derived pairs in $\mathbf{OUT}_4^\perp$. However, for the entailment of a single pair each quasi-reflexive countermodel can be trivially transformed into a reflexive one, so we can use the same embedding into **T**.

Lemma 20. *The structural rules (IO-Wk), (IO-Ctr) and (IO-Cut) are admissible in the calculus $\mathcal{SC}_{\text{OUT}_4^\perp}$.*

Proof. The proof proceeds in the same way as the corresponding proof for $\mathcal{SC}_{\text{OUT}_2^\perp}$ (Lem. 18). The only difference comes from the modified separation principles: outputs can now also appear on the left side of the input sequents. However, this change does not affect any of the case analyses. More precisely, in **Step 1** the formulas $\{X_j\}_{j \in J''}$ will additionally appear on the left in the input sequent (1), and in **Step 2** $\{W_j\}_{j \in J'}$ will additionally appear on the left in the input sequent, but the application of cut to both of them still gives exactly the required input sequent (InS) (which also has all these formulas on the left due to the separation principle). Also $\{X_j\}_{j \in J''}$ additionally appear on the left in the input sequent in **Step 3**; from this sequent, the required input sequent (InS) is obtainable by weakening in **LK**. $\square$

To simplify the completeness proof, we can observe that all derivations in $\mathcal{SC}_{\text{OUT}_2^\perp}$ go in $\mathcal{SC}_{\text{OUT}_4^\perp}$ too.

Lemma 21. *If $G \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\text{OUT}_2^\perp}$ it is also derivable in $\mathcal{SC}_{\text{OUT}_4^\perp}$.*

Proof. By an easy inspection of (E_2) and (E_4) . $\square$

Theorem 26 (Soundness and Completeness of $\mathcal{SC}_{\text{OUT}_4^\perp}$). *$G \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\text{OUT}_4^\perp}$ iff $G \vdash_{\text{OUT}_4^\perp} (B, Y)$.*

Proof. (Completeness). Analogous to the proof of Th. 25. Due to Lem. 21 we only need to check the derivability of the additional principle (CT) in $\mathcal{SC}_{\text{OUT}_4^\perp}$. We first derive $(A, X), (A \wedge X, Y) \vdash (A, Y)$ as follows:

$$\frac{\frac{A \wedge \neg(A \wedge X) \wedge \neg A \vdash}{\vdash (A \wedge \neg(A \wedge X)) \wedge \neg A, Y} \text{(IN)} \quad \frac{A \wedge \neg(A \wedge X) \wedge X \vdash}{\vdash (A \wedge \neg(A \wedge X)) \wedge X, Y \vee \neg X} \text{(IN)} \quad \frac{\vdash Y \vee \neg Y}{(A, X) \vdash (A \wedge Y, Y \vee \neg Y)} \text{(OUT)} \quad \frac{}{(A, X), (A \wedge X, Y) \vdash (A, Y)} \text{(E}_4\text{)}$$

We then apply the structural rule (IO-Cut) with $G \vdash (A, X)$ and then (IO-Cut) with $G \vdash (A \wedge X, Y)$, followed by (IO-Ctr). The claim follows by Lem. 20.

(Soundness). Analogous to the proof of Th. 24. The soundness of the rule (E_4) in $\mathcal{SC}_{\text{OUT}_4^\perp}$ is shown below

$$\frac{\frac{(A, X)}{(A, Y \vee X)} \text{(WO)} \quad \frac{(A, X)}{(B \wedge A, X)} \text{(SI)} \quad \frac{(B \wedge X, Y \vee \neg X)}{(B \wedge A \wedge X, Y \vee \neg X)} \text{(SI)} \quad \frac{}{(B \wedge A, Y \vee \neg X)} \text{(CT)} \quad \frac{}{(B \wedge A, Y \vee X)} \text{(AND)} \quad \frac{(B \wedge A, (Y \vee X) \wedge (Y \vee \neg X))}{(B \wedge A, Y \vee (X \wedge \neg X))} \text{(WO)} \quad \frac{}{(B \wedge A, Y)} \text{(WO)} \quad \frac{(B \wedge \neg A, Y)}{(B \wedge \neg A, Y)} \text{(OR)} \quad \frac{((B \wedge A) \vee (B \wedge \neg A), Y)}{(B \wedge (A \vee \neg A), Y)} \text{(SI)} \quad \frac{}{(B, Y)} \text{(SI)}$$

□

Sequent calculi for $\text{OUT}_1^\perp$ and $\text{OUT}_3^\perp$

The calculi $\mathcal{SC}_{\text{OUT}_1^\perp}$ and $\mathcal{SC}_{\text{OUT}_3^\perp}$ for $\text{OUT}_1^\perp$ and $\text{OUT}_3^\perp$ consist of the base calculus (i.e., **LK** with (IN) and (OUT)) extended with the pair elimination rules (E_1) and (E_3) in Fig. 5.1, respectively.

Remark 7. *Unlike $\text{OUT}_2^\perp$ and $\text{OUT}_4^\perp$, there is no modal embedding in the literature to provide guidance for the development of the rules for $\text{OUT}_1^\perp$ and $\text{OUT}_3^\perp$. A helpful way to understand the new rules is to view them as restricted versions of the pair-elimination rules $\text{OUT}_2^\perp$ and $\text{OUT}_4^\perp$. The key difference is that $\text{OUT}_1^\perp$ and $\text{OUT}_3^\perp$ forbid us from freely combining inputs. In the stronger systems, the elimination rules allow us to extend the goal input with whatever input is carried by the deriving formula. But in $\text{OUT}_1^\perp$ and $\text{OUT}_3^\perp$ such combinations are not always allowed. To respect this limitation, the new rules replace the “input-combining” premise with an **LK** sequent that explicitly requires the deriving input to be provable from the goal input. In other words, any use of an additional input must now be justified by a separate derivation: the systems no longer grants it for free. The rules (E_1) and (E_3) are defined by reflecting this.*

Because the rules $\mathcal{SC}_{\text{OUT}_1^\perp}$ and $\mathcal{SC}_{\text{OUT}_3^\perp}$ include an explicit **LK** premise, their derivations are simpler than those in $\mathcal{SC}_{\text{OUT}_2^\perp}$ and $\mathcal{SC}_{\text{OUT}_4^\perp}$. Although this structure could in principle be exploited directly in the soundness and completeness proofs, we follow the same strategy used for the latter calculi and instead establish appropriate separation principles. These principles will play a central role in settling open questions about computational bounds and modal embeddings for $\text{OUT}_1^\perp$ and $\text{OUT}_3^\perp$. The proof of the separation lemma for $\mathcal{SC}_{\text{OUT}_1^\perp}$ and $\mathcal{SC}_{\text{OUT}_3^\perp}$ is less direct than for the previous cases. The underlying intuition is that the characterization inspects all possible ways of applying the rule (E_3) (or (E_1)), by partitioning the deriving pairs $(A_1, X_1), \dots, (A_n, X_n)$ into two disjoint sets I of remaining deriving pairs and J of eliminated pairs). This decomposition yields a characterization of derivability of I/O sequents in terms of the derivability of *three* **LK** sequents.

We present the proof of the lemma for $\mathcal{SC}_{\text{OUT}_3^\perp}$, the one for $\mathcal{SC}_{\text{OUT}_1^\perp}$ being a simplified case.

Lemma 22 (Separation principle for $\mathcal{SC}_{\text{OUT}_3^\perp}$). *$(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\text{OUT}_3^\perp}$ iff for all $(I, J) \in \mathcal{P}(\{1, \dots, n\})$, one of the following holds:*

- $B, \{X_j\}_{j \in J} \vdash A_i$ (input sequent) is derivable in **LK** for some $i \in I$,
- $B, \{X_j\}_{j \in J} \vdash (\text{ex-falso sequent})$ is derivable in **LK**,
- $\{X_j\}_{j \in J} \vdash Y$ (output sequent) is derivable in **LK**.

Proof. We follow the general proof strategy used for Lem. 17, traversing the given derivation bottom-up, now following only the right premise (as this is the only I/O sequent in (E_3)).

(Left-to-right). Suppose towards contradiction that $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$, and fix a partition $(I, J) \in \mathcal{P}(\{1, \dots, n\})$ for which all the input sequents, the ex-falso sequent and the output sequent are underivable in $\mathbf{LK}$. We now trace the given derivation bottom-up through applications of the rule (E_3) . At each application we proceed to the right premise and record the index of the eliminated pair into a growing set J' (starting with $J' = \emptyset$). At each step the goal pair has the form $(B \wedge \bigwedge_{j \in J'} X_j, Y \vee \bigvee_{j \in J'} \neg X_j)$. Notice that every index k added to J' belongs to J , not to I . Indeed, the left premise of the rule (E_3) at that step is $B \wedge \bigwedge_{j \in J'} X_j \vdash A_k$ and if $k \in I$ the derivability of such sequent would imply the derivability of one of the input sequents due to the admissibility of weakening in $\mathbf{LK}$ (assuming inductively that $J' \subseteq J$ on the previous step), contrary to the assumption. Eventually, we reach the I/O sequent to which one of the transition rules is applied. If this rule is (IN) then the derivability of its premise implies derivability of the ex-falso sequent in $\mathbf{LK}$. If it is (OUT) then the derivability of its premise implies derivability of the output sequent in $\mathbf{LK}$. Either case contradicts our initial assumption that these sequents are underivable.

(Right-to-left). Conversely, suppose that the sequent $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is not derivable. We apply the pair elimination rule to it iteratively, obtaining new non-derivable I/O sequents, and construct a partition that violates the separation principle. Specifically, we iteratively apply the rule (E_3) . At each step, if the left premise of (E_3) eliminating some pair on the left is derivable in $\mathbf{LK}$, we eliminate that pair and move to the right premise. Since the right premise must be non-derivable (as so is assumed to be the sequent $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$) in $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$, this process always continues through right premises. The number of deriving pairs strictly decreases at each step, so after finitely many iterations no further such rule application is possible. Let J be the set of indices of the pairs eliminated in previous iterations, and I the set of indices of the pairs remaining on the left. The pair on the right has the form $(B \wedge \bigwedge_{j \in J} X_j, Y \vee \bigvee_{j \in J} \neg X_j)$. The partition (I, J) violates the separation principle. Indeed:

- each input sequent for $i \in I$ is underivable; otherwise we could eliminate (A_i, X_i) at the previous step, contrary to the construction;
- the ex-falso sequent is underivable, otherwise the final sequent would be derivable in $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$ via (IN) ;
- the output sequent is underivable, otherwise the final sequent would be derivable via (OUT) .

Therefore (I, J) witnesses the failure of the separation principle for the given non-derivable sequent. $\square$

The separation principle for $\mathcal{SC}_{\mathbf{OUT}_1^\perp}$ has the following formulation (the proof is the same as the proof for $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$).

Lemma 23 (Separation principle for $\mathcal{SC}_{\mathbf{OUT}_1^\perp}$). *$(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\mathbf{OUT}_1^\perp}$ iff for all partitions $(I, J) \in \mathcal{P}(\{1, \dots, n\})$, at least one of the following holds:*

- $B \vdash A_i$ (input sequent) is derivable in LK for some $i \in I$,
- $B \vdash$ (ex-falso sequent) is derivable in LK ,
- $\{X_j\}_{j \in J} \vdash Y$ (output sequent) is derivable in LK .

With these separation principles, the soundness and completeness proofs can be done the same way as before.

Lemma 24. *The rules (IO-Wk), (IO-Ctr) and (IO-Cut) are admissible for the calculi $\mathcal{SC}_{\mathbf{OUT}_1^\perp}$ and $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$.*

Proof. Easy for (IO-Wk) and (IO-Ctr). For cut admissibility, we follow the same three-step strategy used in the proof of Lem. 18. The only difference is that, for $\mathcal{SC}_{\mathbf{OUT}_1^\perp}$ and $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$, the separation principles include the ex-falso sequent and potentially several input sequents. This increases the number of cases to check at each step, but all of them are handled in the same way as before. We describe the argument for $\mathcal{SC}_{\mathbf{OUT}_1^\perp}$ as it is closest to the proof given in full for Lem. 18. Instead of deriving just the two sequents (InS) and $(OutS)$, we must now obtain, one of the following:

- the output sequent $(OutS) \{X_j\}_{j \in J''}, \{W_j\}_{j \in J'} \vdash Y$,
- the ex-falso sequent $(ExFalsoS) B \vdash$,
- one of the input sequents $(InS''_i) B \vdash A_i$ for some $i \in I''$ or one of the input sequents $(InS'_i) B \vdash D_i$ for some $i \in I'$.

Step 1: all cases except the input sequent (1) $B \vdash A_0$ immediately yield one of the required sequents by weakening in **LK**. **Step 2:** if any input sequent $A_0 \vdash D_i$ is derivable, then cutting it with (1) gives exactly the required sequent (InS'_i) ; the same applies to the ex-falso sequent $A_0 \vdash$, as cutting it with (1) yields $(ExFalsoS)$. The only remaining case is (2) $\{W_j\}_{j \in J'} \vdash X_0$. Finally, in **Step 3**, the cases of all input sequents and ex-falso sequent are again handled by weakening in **LK**, and the case of the output sequent is again handled by cut in **LK** with (2) yielding $(OutS)$. Reasoning for $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$ is exactly the same, the output formulas appearing on the left do not interfere with obtaining the required sequents by either weakening or cut in **LK**. $\square$

Theorem 27 (Soundness and completeness of $\mathcal{SC}_{\mathbf{OUT}_1^\perp}$ and $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$). For $k \in \{1, 3\}$, $G \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\mathbf{OUT}_k^\perp}$ iff $G \vdash_{\mathbf{OUT}_k^\perp} (B, Y)$.

Proof. The proof proceeds as for $\mathcal{SC}_{\mathbf{OUT}_2^\perp}$ and $\mathcal{SC}_{\mathbf{OUT}_4^\perp}$.

(Completeness) Assume that (B, Y) is derivable in $\mathbf{OUT}_1^\perp$ (resp. $\mathbf{OUT}_3^\perp$). We prove by induction on the derivation tree that for each pair (A, X) occurring in it, the I/O sequent $G \vdash (A, X)$ is derivable in $\mathcal{SC}_{\mathbf{OUT}_1^\perp}$ (resp. $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$). The case $(A, X) \in G$ is now handled as follows:

$$\frac{A \vdash A \quad \frac{\vdash X \vee \neg X}{G' \vdash (A, X \vee \neg X)} (\text{OUT})}{(A, X), G' \vdash (A, X)} (E_1)$$

The simulations of the original rules (WO) and (SI) is as follows:

$$\frac{\frac{\frac{X \vdash Y}{\vdash Y, \neg X} (\neg R) \quad \frac{\vdash Y \vee \neg X}{\vdash (A, Y \vee \neg X)} (\vee R)}{A \vdash A \quad \vdash (A, Y \vee \neg X)} (\text{OUT})}{(A, X) \vdash (A, Y)} (E_1) \quad \frac{B \vdash A \quad \frac{\vdash Y \vee \neg Y}{\vdash (B, Y \vee \neg Y)} (\text{OUT})}{(A, Y) \vdash (B, Y)} (E_1)$$

The derivation of the sequent for rule (AND) is as follows (rule (AND) is then simulated by applying two cuts and a series of contractions to it as before):

$$\frac{B \vdash B \quad \frac{\vdash (X_1 \wedge X_2) \vee \neg X_1 \vee \neg X_2}{\vdash (B, (X_1 \wedge X_2) \vee \neg X_1 \vee \neg X_2)} (\text{OUT})}{(B, X_2) \vdash (B, (X_1 \wedge X_2) \vee \neg X_1)} (E_1) \quad \frac{B \vdash B \quad (B, X_2) \vdash (B, (X_1 \wedge X_2) \vee \neg X_1)}{(B, X_1), (B, X_2) \vdash (B, X_1 \wedge X_2)} (E_1)$$

The derivations of (TOP) and (BOT) is immediate.

For the completeness of $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$, first observe that, similarly to Lem. 21, we can prove that every sequent derivable in $\mathcal{SC}_{\mathbf{OUT}_1^\perp}$ is also derivable in $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$. A derivation of the additional principle (CT) in $\mathbf{OUT}_3^\perp$ is as follows

$$\frac{A \wedge X \vdash A \wedge X \quad \frac{\vdash Y \vee \neg X \vee \neg Y}{\vdash (A \wedge X \wedge Y, Y \vee \neg X \vee \neg Y)} (\text{OUT})}{\frac{A \wedge X \vdash A \wedge X \quad (A \wedge X, Y) \vdash (A \wedge X, Y \vee \neg X)}{(A, X), (A \wedge X, Y) \vdash (A, Y)} (E_3)} (E_3)$$

The claim follows by the admissibility proof of the structural rules in $\mathcal{SC}_{\mathbf{OUT}_1^\perp}$ and $\mathcal{SC}_{\mathbf{OUT}_3^\perp}$ (Lem. 24).

(Soundness): To establish the soundness of the rule (E_1) in $\mathbf{OUT}_1^\perp$, observe that if the premise $B \vdash A$ is derivable in **LK**, then $A \in \text{Cn}(\{B\})$, and hence $(B \wedge A) \in \text{Cn}(\{B\})$.

The derivation below, demonstrating the soundness for rule (E_1) in $\mathbf{OUT}_1^\perp$, uses this fact at the last step.

$$\begin{array}{c}
 \frac{(A, X)}{(A, Y \vee X)} \text{ (WO)} \\
 \frac{(A, Y \vee X)}{(B \wedge A, Y \vee X)} \text{ (SI)} \quad \frac{(B, Y \vee \neg X)}{(B \wedge A, Y \vee \neg X)} \text{ (SI)} \\
 \frac{(B \wedge A, Y \vee X) \quad (B \wedge A, Y \vee \neg X)}{(B \wedge A, (Y \vee X) \wedge (Y \vee \neg X))} \text{ (AND)} \\
 \frac{(B \wedge A, (Y \vee X) \wedge (Y \vee \neg X))}{(B \wedge A, Y \vee (X \wedge \neg X))} \text{ (WO)} \\
 \frac{(B \wedge A, Y \vee (X \wedge \neg X))}{(B \wedge A, Y)} \text{ (WO)} \\
 \frac{(B \wedge A, Y)}{(B, Y)} \text{ (SI)}
 \end{array}$$

For the soundness of the rule (E_3) in $\mathbf{OUT}_3^\perp$ we have the following analagous derivation:

$$\begin{array}{c}
 \frac{(A, X)}{(A, Y \vee X)} \text{ (WO)} \quad \frac{(A, X)}{(B \wedge A, X)} \text{ (SI)} \quad \frac{(B \wedge X, Y \vee \neg X)}{(B \wedge A \wedge X, Y \vee \neg X)} \text{ (SI)} \\
 \frac{(A, Y \vee X)}{(B \wedge A, Y \vee X)} \text{ (SI)} \quad \frac{(B \wedge A, X) \quad (B \wedge A \wedge X, Y \vee \neg X)}{(B \wedge A, Y \vee \neg X)} \text{ (CT)} \\
 \frac{(B \wedge A, Y \vee X) \quad (B \wedge A, Y \vee \neg X)}{(B \wedge A, (Y \vee X) \wedge (Y \vee \neg X))} \text{ (AND)} \\
 \frac{(B \wedge A, (Y \vee X) \wedge (Y \vee \neg X))}{(B \wedge A, Y \vee (X \wedge \neg X))} \text{ (WO)} \\
 \frac{(B \wedge A, Y \vee (X \wedge \neg X))}{(B \wedge A, Y)} \text{ (WO)} \\
 \frac{(B \wedge A, Y)}{(B, Y)} \text{ (SI)}
 \end{array}$$

□

5.4.2 Sequent Calculi for Original I/O logics

Original (deontic) I/O logics differ from their causal variants by the absence of the principle (BOT). In the calculi for the latter logics presented above, this principle is tightly connected with the transition rule (IN) . While for the systems without (OR) the remaining rules already form a sound and complete calculus, for the systems with (OR) additional modifications are required. We start with the former case.

Sequent calculi for $\mathbf{OUT}_1$ and $\mathbf{OUT}_3$

For these logics, the principle (BOT) is the only source of application of the transition rule (IN) in the corresponding calculi, and conversely (BOT) appears only as part of the simulation of (IN) . Thus, once (IN) is removed, the soundness and completeness proofs go through unchanged, yielding calculi for the original versions of these logics.

These logics also admit separation principles analogous to those of their causal counterparts, differing only for the absence of the ex-falso sequent: in $\mathbf{OUT}_1^\perp$ and $\mathbf{OUT}_3^\perp$ this sequent arises solely from the rule (IN) , hence the separation principles for $\mathbf{OUT}_1$ and $\mathbf{OUT}_3$ are simply obtained by dropping it.

Lemma 25 (Separation principles for $\mathbf{OUT}_1$ and $\mathbf{OUT}_3$). *The following two equivalences hold.*

- $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in $SC_{\mathbf{OUT}_1}$ iff for every partition $(I, J) \in \mathcal{P}(\{1, \dots, n\})$, either $\{X_j\}_{j \in J} \vdash Y$ or $B \vdash A_i$ for some $i \in I$ is derivable in $\mathbf{LK}$.

- $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in $SC_{\mathbf{OUT}_3}$ iff for every partition $(I, J) \in \mathcal{P}(\{1, \dots, n\})$, either $\{X_j\}_{j \in J} \vdash Y$ or $B, \{X_j\}_{j \in J} \vdash A_i$ for some $i \in I$ is derivable in $\mathbf{LK}$.

Proof. Analogous to the proofs of Lem. 22 and 23. The only difference is the absence of the case of the transition rule (IN) , corresponding in both directions to the ex-falso sequent, which therefore is not present in the lemma's statement. $\square$

The proof of the admissibility of the structural rules can be easily adjusted.

Lemma 26. *The rules $(IO-Wk)$, $(IO-Ctr)$, and $(IO-Cut)$ are admissible in $SC_{\mathbf{OUT}_1}$ and $SC_{\mathbf{OUT}_3}$.*

Proof. Same as the proof of Lem. 24. In the cut-admissibility proof, the ex-falso sequent no longer arises in any step, and thus no case is handled by deriving the sequent $(ExFalsoS)$. $\square$

The soundness and completeness proofs remain the same, with one less case to consider in both directions, due to the absence of (BOT) (and of (IN)).

Theorem 28 (Soundness and completeness of the calculi for $\mathbf{OUT}_1$ and $\mathbf{OUT}_3$). *For both $k \in \{1, 3\}$, $G \vdash (B, Y)$ is derivable in $SC_{\mathbf{OUT}_k}$ iff $G \vdash_{\mathbf{OUT}_k} (B, Y)$.*

Sequent calculi for $\mathbf{OUT}_2$ and $\mathbf{OUT}_4$

Calculi for $\mathbf{OUT}_2$ and $\mathbf{OUT}_4$ cannot be obtained by simply removing the rule (IN) from $SC_{\mathbf{OUT}_2^\perp}$ and $SC_{\mathbf{OUT}_4^\perp}$, since this rule is used not only in the derivation of (BOT) , but also throughout the derivations of almost all other I/O principles (see the proofs of Th. 25 and Th. 26). Sound and complete calculi for $\mathbf{OUT}_2$ and $\mathbf{OUT}_4$ are obtained by replacing in $SC_{\mathbf{OUT}_2^\perp}$ (resp. $SC_{\mathbf{OUT}_4^\perp}$) the rule (IN) with the rule (E_1) (resp. (E_3)), and also explicitly adding $(IO-Ctr)$. We call the resulting calculi $SC_{\mathbf{OUT}_2}$ and $SC_{\mathbf{OUT}_4}$ (Fig. 5.1).

Theorem 29 (Soundness). *If $G \vdash (B, Y)$ is derivable in $SC_{\mathbf{OUT}_2}$ (resp. $SC_{\mathbf{OUT}_4}$) then $G \vdash_{\mathbf{OUT}_2} (B, Y)$ (resp. $G \vdash_{\mathbf{OUT}_4} (B, Y)$).*

Proof. The rules (E_1) and (E_3) are already known to be sound for the original I/O logics, since we have shown that each can be emulated in the strictly weaker I/O logics $\mathbf{OUT}_1$ and $\mathbf{OUT}_3$, respectively (see the proof of Th. 27). $(IO-Ctr)$ is trivially sound. $\square$

The technical lemma below, needed to prove the separation principles, shows that applications of (IN) in left subderivations of (E_2) (resp. (E_4)) can be simulated by (E_1) (resp. (E_3)) and $(IO-Ctr)$.

Lemma 27. *The two following statements hold.*

- *If $G \vdash (B \wedge \neg A, Y)$ is derivable in $SC_{OUT_2^\perp}$ and $G \vdash (B, Y \vee \neg X)$ is derivable in SC_{OUT_2} then $(A, X), G \vdash (B, Y)$ is derivable in SC_{OUT_2} .*
- *If $G \vdash (B \wedge \neg A, Y)$ is derivable in $SC_{OUT_4^\perp}$ and $G \vdash (B \wedge X, Y \vee \neg X)$ is derivable in SC_{OUT_4} then $(A, X), G \vdash (B, Y)$ is derivable in SC_{OUT_4} .*

Proof. We prove the result for SC_{OUT_4} ; the proof for SC_{OUT_2} is analogous and simpler. Let $\mathfrak{D}_1$ be the derivation of $G \vdash (B \wedge \neg A, Y)$ in $SC_{OUT_4^\perp}$, and $\mathfrak{D}_2$ be the derivation of $G \vdash (B \wedge X, Y \vee \neg X)$ in SC_{OUT_4} .

We show how to obtain the desired derivation in SC_{OUT_4} of $(A, X), G \vdash (B, Y)$. Start from $\mathfrak{D}_1$ and uniformly add $G, (A, X)$ to the left of all the I/O sequents. This yields a derivation of $G, (A, X), G \vdash (B \wedge \neg A, Y)$ in $SC_{OUT_4^\perp}$. The only obstacle to viewing this derivation as one in SC_{OUT_4} is the possible presence of applications of the rule (IN) . Each such application concludes a sequent of the form $G, (A, X), G' \vdash (B \wedge \neg A \wedge C, Y \vee D)$ where G' is a subset of G , C (resp. D) is a conjunction (resp. disjunction) of a possibly empty list of formulas, with the **LK** premise $B \wedge \neg A \wedge C \vdash$. We eliminate this use of (IN) as follows. Apply the pair elimination rule (E_3) to (A, X) , yielding the premises $B \wedge \neg A \wedge C \vdash A$ and $G, G' \vdash (B \wedge \neg A \wedge C \wedge X, Y \vee D \vee \neg X)$. The first premise is derivable in **LK** (being equivalent to $B \wedge \neg A \wedge C \vdash$, which is the premise of the original application of (IN)). To obtain a derivation of the second premise in SC_{OUT_4} , take $\mathfrak{D}_2$ and replace each I/O sequent $G'' \vdash (B \wedge X \wedge C'', Y \vee \neg X \vee D'')$ in it with sequent $G, G'' \vdash (B \wedge \neg A \wedge C \wedge X \wedge C'', Y \vee D \vee \neg X \vee D'')$ (note that all applications of the rules (OUT) , (E_3) , and (E_4) in $\mathfrak{D}_2$ remain correct after such transformation). Performing this replacement to all applications of (IN) yields a derivation in SC_{OUT_4} of $G, G, (A, X) \vdash (B \wedge \neg A, Y)$. The desired derivation of $G, (A, X) \vdash (B, Y)$ in SC_{OUT_4} follows by $(IO\text{-}Ctr)$ and (E_4) , using the derivation $\mathfrak{D}_2$ of $G \vdash (B \wedge X, Y \vee \neg X)$. $\square$

The intuition behind the simulation of (IN) is mirrored in the separation principles, which differ from the principles for $SC_{OUT_2^\perp}/SC_{OUT_4^\perp}$ only in that the input sequent for the partition (I, J) with $I = \emptyset$ is omitted. Recall that in the separation principles for $SC_{OUT_2^\perp}/SC_{OUT_4^\perp}$, adding an index i to I represents moving to the left premise when eliminating the pair (A_i, X_i) , and the input sequent corresponds to the application of the rule (IN) .

Lemma 28 (Separation principle for SC_{OUT_4}). *$(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in SC_{OUT_4} iff for all partitions $(I, J) \in \mathcal{P}(\{1, \dots, n\})$, either $\{X_j\}_{j \in J} \vdash Y$ (output sequent) is derivable in **LK** or both $I \neq \emptyset$ and $B, \{X_j\}_{j \in J} \vdash \{A_i\}_{i \in I}$ (input sequent) is derivable in **LK**.*

Proof. We follow the general proof strategy for separation principles (see the proof of Lem. 17), adapting both the argument and the rule-application policy to the presence

of two pair-elimination rules. In effect, the proof combines the proofs of the separation principles for $\mathcal{SC}_{\text{OUT}_4^\perp}$ (Lem. 19) and $\mathcal{SC}_{\text{OUT}_3^\perp}$ (Lem. 22).

(Left-to-right). Suppose towards contradiction the existence of a derivation for sequent $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ and a partition $(I, J) \in \mathcal{P}(\{1, \dots, n\})$ such that the output sequent in the separation principle is underivable and, moreover, either $I = \emptyset$ or the corresponding input sequent is also underivable. We traverse the derivation bottom-up, following applications of (E_3) and (E_4) . When eliminating a pair (A_i, X_i) via (E_4) we move to the left premise if $i \in I$ and to the right premise if $i \in J$; when eliminating via (E_3) we always move to the right premise. Let I' (resp. J') be the set of indices of eliminated pairs for which we moved to the left (resp. right) premise. As in the proof of Lem. 22, every use of (E_3) must eliminate a pair with index in J (otherwise $I \neq \emptyset$ and the input sequent can be derived in $\mathbf{LK}$ via weakening from the left premise of this application of (E_3) , contradicting the initial assumption); hence $I' \subseteq I$ and $J' \subseteq J$. Eventually, we reach the transition rule (OUT) applied to the goal pair $(B \wedge \bigwedge_{i \in I'} \neg A_i \wedge \bigwedge_{j \in J'} X_j, Y \wedge \bigvee_{j \in J'} \neg X_j)$. Thus the sequent $\vdash Y \wedge \bigvee_{j \in J'} \neg X_j$ is derivable in $\mathbf{LK}$, and by weakening so is the required output sequent, contradicting our initial assumption.

(Right-to-left). We follow the same iterative construction as in the proof of Lem. 17, except that instead of applying rule (E_4) , we use Lem. 27 in its contrapositive form. Specifically, if $(A, X), G \vdash (B, Y)$ is not derivable in $\mathcal{SC}_{\text{OUT}_4}$, then either $G \vdash (B \wedge \neg A, Y)$ is not derivable in $\mathcal{SC}_{\text{OUT}_4^\perp}$ or $G \vdash (B, Y \vee \neg X)$ is not derivable in $\mathcal{SC}_{\text{OUT}_4}$. We apply this fact iteratively, moving to the underivable premise, and if we move to the left premise, we switch to non-derivability in $\mathcal{SC}_{\text{OUT}_4^\perp}$ and proceed as in the proof of Lem. 19 applying rule (E_4) . After removing all deriving pairs, if in the end $I \neq \emptyset$ we have non-derivability in $\mathcal{SC}_{\text{OUT}_4^\perp}$ and therefore can apply both (IN) and (OUT) obtaining non-derivability of both the input and the output sequent, while if $I = \emptyset$ we have non-derivability in $\mathcal{SC}_{\text{OUT}_4}$ and can only apply (OUT) obtaining non-derivability of the output sequent. This matches exactly the separation principle for $\mathcal{SC}_{\text{OUT}_4}$. $\square$

The proof for $\mathcal{SC}_{\text{OUT}_2}$ is a simplified version of the proof for $\mathcal{SC}_{\text{OUT}_4}$.

Lemma 29 (Separation principle for $\mathcal{SC}_{\text{OUT}_2}$). *$(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\text{OUT}_2}$ iff for all partitions $(I, J) \in \mathcal{P}(\{1, \dots, n\})$, either $\{X_j\}_{j \in J} \vdash Y$ is derivable in $\mathbf{LK}$ or both $I \neq \emptyset$ and $B \vdash \{A_i\}_{i \in I}$ is derivable in $\mathbf{LK}$.*

Example 11. *Following [BPvdT20], consider the following sentences from the European Data Protection Regulation (GDPR, Regulation EU 2016/679):*

- (1) *Personal data shall be processed lawfully (Art. 5),*
- (2) *If the personal data have been processed unlawfully, the controller has the obligation to erase the personal data in question without delay (Art. 17.d, right to be forgotten),*

- (3) *It is obligatory, e.g., as part of a respective agreement between a customer and a company, to keep the personal data (as relevant to the agreement) provided that it is processed lawfully.*

Their formalization in (the original) I/O logics consists of conditional norms

$$G = \{(\top, \text{Lawful}), (\neg \text{Lawful}, \text{Erase}), (\text{Lawful}, \neg \text{Erase})\},$$

where *Lawful* is a propositional atom standing for ‘data is processed lawfully’ and *Erase* is a propositional atom standing for ‘data is erased’. The sentences (1)-(3) have a contrary-to-duty structure (that is, they contain obligations that, once violated, trigger further obligations), and when formalized in a naive deontic logic like SDL [vW51], they allow the derivation of any other obligation (deontic explosion). This is not the case in the original I/O logics; we use the separation principle to show that an unethical obligation under condition $\neg \text{Lawful}$ (like $(\neg \text{Lawful}, \text{KillMary})$, as proposed in [BPvdT20]) cannot be derived in **OUT₂**. Indeed, let $(\top, \text{Lawful})$ be (A_1, X_1) , $(\neg \text{Lawful}, \text{Erase})$ be (A_2, X_2) , and $(\text{Lawful}, \neg \text{Erase})$ be (A_3, X_3) . A counterexample to the entailment $G \vdash_{\text{OUT}_2} (\neg \text{Lawful}, \text{KillMary})$ is provided by the partition $(\{3\}, \{1, 2\})$, for which neither the output sequent

$$\text{Lawful}, \text{Erase} \vdash \text{KillMary},$$

nor the input sequent

$$\text{Lawful} \vdash \neg \text{Lawful}$$

is classically valid. This partition however is not a separating partition for this entailment in **OUT₄**, and in fact this entailment holds in **OUT₄** (and **OUT₃**), as demonstrated by the following derivation in **SC_{OUT₄}** (we use abbreviated versions of propositional atoms):

$$\frac{\neg L \vdash \top \quad \frac{\neg L \wedge L \vdash \neg L \quad \frac{(L, \neg E) \vdash (\neg L \wedge L \wedge E, KM \vee \neg L \vee \neg E)}{(E_3)} \quad \frac{\neg L \wedge L \wedge E \vdash L \quad \frac{\vdash KM \vee \neg L \vee \neg E \vee \neg \neg E}{(OUT)} \quad \frac{\neg L \wedge L \wedge E \wedge \neg E, KM \vee \neg L \vee \neg E \vee \neg \neg E}{(E_3)}}{(E_3)} \quad \frac{(\neg L, E), (L, \neg E) \vdash (\neg L \wedge L, KM \vee \neg L)}{(E_3)}}{(\top, L), (\neg L, E), (L, \neg E) \vdash (\neg L, KM)}$$

The separation principles can be used to prove the admissibility of the structural rules (IO-Wk), and (IO-Cut) in **SC_{OUT₂}** and **SC_{OUT₄}** analogously to their causal counterparts.

Lemma 30. *The structural rules (IO-Wk) and (IO-Cut) are admissible in **SC_{OUT₂}** and **SC_{OUT₄}**.*

Proof. Analogous to the proofs of Lem. 18 and Lem. 20. In the cut admissibility proof, we distinguish two cases. First, when $I' \cup I'' \neq \emptyset$ the same proof can be used: by the modified separation principles we are still allowed to derive either the input sequent

(*InS*) or the output sequent (*OutS*) and there are just less cases to consider if either $I' = \emptyset$ or $I'' = \emptyset$. Second, when $I' \cup I'' = \emptyset$ we cannot derive the input sequent (*InS*) due to the modified separation principle. Nevertheless, the same proof strategy applies: since in this situation both $I = \emptyset$ and $I'' = \emptyset$ and therefore we do not need to consider the cases of the input sequent in Step 2 and Step 3. It is easy to check that all the remaining cases conclude with the derivation of (*OutS*). $\square$

At the same time, (IO-Ctr) is not admissible in these calculi. A counterexample for $\mathcal{SC}_{\mathbf{OUT}_4}$ is given below. (A similar example can also be constructed for $\mathcal{SC}_{\mathbf{OUT}_2}$).

Example 12. *The sequent $(\neg X, X), (A \wedge X, Y), (A \wedge Y, \neg X) \vdash (A, X \wedge Y)$ can be derived in $\mathcal{SC}_{\mathbf{OUT}_4}$ as follows, using the rule (IO-Ctr) to introduce a copy of the pair $(\neg X, X)$ and then immediately eliminate it via (E_4):*

$$\frac{\frac{\spadesuit_1}{(\neg X, X), (\neg X, X), (A \wedge X, Y), (A \wedge Y, \neg X) \vdash (A, X \wedge Y)} \quad \frac{\spadesuit_2}{(\neg X, X), (A \wedge X, Y), (A \wedge Y, \neg X) \vdash (A, X \wedge Y)} (E_4)}{(\neg X, X), (A \wedge X, Y), (A \wedge Y, \neg X) \vdash (A, X \wedge Y)} (IO-Ctr)$$

with the following subderivation $\spadesuit_1$ on the left:

$$\frac{A \wedge X \vdash A \wedge X \quad \frac{A \wedge X \wedge Y \vdash A \wedge Y \quad \frac{A \wedge X \wedge Y \wedge \neg X \vdash \neg X \quad \frac{\vdash (X \wedge Y) \vee \neg Y \vee X \vee \neg X}{\vdash (A \wedge X \wedge Y \wedge \neg X \wedge X, (X \wedge Y) \vee \neg Y \vee X \vee \neg X)} (OUT)}{(\neg X, X) \vdash (A \wedge X \wedge Y \wedge \neg X, (X \wedge Y) \vee \neg Y \vee X)} (E_3)}{(\neg X, X), (A \wedge Y, \neg X) \vdash (A \wedge X \wedge Y, (X \wedge Y) \vee \neg Y)} (E_3)}{(\neg X, X), (A \wedge X, Y), (A \wedge Y, \neg X) \vdash (A \wedge X, X \wedge Y)} (E_3)$$

and the following subderivation $\spadesuit_2$ on the right:

$$\frac{A \wedge X \vdash A \wedge X \quad \frac{\vdash (X \wedge Y) \vee \neg X \vee \neg Y}{(\neg X, X), (A \wedge Y, \neg X) \vdash (A \wedge X \wedge Y, (X \wedge Y) \vee \neg X \vee \neg Y)} (OUT)}{(\neg X, X), (A \wedge X, Y), (A \wedge Y, \neg X) \vdash (A \wedge X, (X \wedge Y) \vee \neg X)} (E_3)$$

It is not hard to see that no derivation can be found without duplicating the deriving pairs. Indeed, intuitively, to successfully conclude the derivation, we have to reach goal outputs containing the disjunction $(X \wedge Y) \vee \neg X \vee \neg Y$. Yet, independently of the deriving pair we start with, these formulas cannot be obtained using the pair $(\neg X, X)$ (resp. $(A \wedge X, Y)$) only once. More precisely, at least one of these pairs must first be eliminated with (E_4) (as the **LK** premise of (E_3) would not be derivable) and then again with (E_3) as the goal output of the left premise of (E_4) does not carry $\neg X$ (resp. $\neg Y$).

We therefore have to add (IO-Ctr) to the calculi $\mathcal{SC}_{\mathbf{OUT}_2}$ and $\mathcal{SC}_{\mathbf{OUT}_4}$ explicitly. We can now prove completeness of these calculi.

Theorem 30 (Completeness of $\mathcal{SC}_{\mathbf{OUT}_2}$ and $\mathcal{SC}_{\mathbf{OUT}_4}$). *If (B, Y) is entailed by G in $\mathbf{OUT}_2$ (resp. $\mathbf{OUT}_4$) then $G \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\mathbf{OUT}_2}$ (resp. $\mathcal{SC}_{\mathbf{OUT}_4}$).*

Proof. Follows the general completeness proof strategy in Th. 25. We only need to provide a derivation of (OR) in the new calculi, since we have already established derivability of the principle (CT) in $\mathcal{SC}_{OUT_3}$ (the rules of which are present in $\mathcal{SC}_{OUT_4}$), and all other principles of the original I/O logics in both $\mathcal{SC}_{OUT_1}$ and $\mathcal{SC}_{OUT_3}$ (the rules of which are present in $\mathcal{SC}_{OUT_2}$ and $\mathcal{SC}_{OUT_4}$ respectively).

The derivation of the sequent for (OR) in $\mathcal{SC}_{OUT_2}$ is as follows (the derivation in $\mathcal{SC}_{OUT_4}$ is analogous):

$$\frac{\frac{(A_1 \vee A_2) \wedge \neg A_1 \vdash A_2 \quad \frac{\vdash Y \vee \neg Y}{\vdash ((A_1 \vee A_2) \wedge \neg A_1, Y \vee \neg Y)} (OUT)}{(A_2, Y) \vdash ((A_1 \vee A_2) \wedge \neg A_1, Y)} (E_1) \quad \frac{\vdash Y \vee \neg Y}{(A_2, Y) \vdash (A_1 \vee A_2, Y \vee \neg Y)} (OUT)}{(A_1, Y), (A_2, Y) \vdash (A_1 \vee A_2, Y)} (E_2)$$

□

Although the calculi $\mathcal{SC}_{OUT_2}$ and $\mathcal{SC}_{OUT_4}$ require contraction, their applications can be bounded (and, accordingly, can still be considered analytic). As shown below, it suffices to consider at most two copies of each deriving pair.

Lemma 31. *If $G \vdash (B, Y)$ is derivable in $\mathcal{SC}_{OUT_2}$ (resp. $\mathcal{SC}_{OUT_4}$) then so is $G, G \vdash (B, Y)$ without using (IO-Ctr).*

Proof. Assume that, in a given branch of the derivation of $G, G \vdash (B, Y)$, a pair (A, X) is used more than twice as a premise of an elimination rule. There are two possibilities. **Case 1:** (A, X) is eliminated twice using (E_2) (resp. (E_4)). Then in the second (higher) application (read bottom-up) one of its premises coincides with the goal pair occurring in the rule's conclusion by idempotency of $\wedge$ and $\vee$. Hence the second elimination does not introduce any new information, and therefore this application of (E_2) (resp. (E_4)) can be removed without affecting derivability. **Case 2:** (A, X) is eliminated twice using (E_1) (resp. (E_3)). By the idempotency of the connectives the higher application of the rule leads to a premise which coincides with the rule's conclusion. Hence, the second elimination can again be removed.

By systematically performing this simplification on every branch of the derivation, we obtain a derivation of $G, G \vdash (B, Y)$ in which no pair is eliminated more than once by the same rule. Hence, applications of (IO-Ctr) are redundant, and can be removed. □

5.4.3 Sequent Calculi for Throughput I/O logics

In throughput logics, the principle (ID) allows derivations to start from any pair (A, A) , and after one application of (WO) , from any pair (A, B) with $B \in Cn(A)$. To obtain the sequent calculi $\mathcal{SC}_{OUT_i^+}$ for these logics, it suffices to replace the transition rules of the corresponding causal calculi with the rule $(THROU)$ from Fig. 5.1, which end the I/O phase of the derivation whenever the goal pair is of this form.

The separation principles for the throughput logics arise from straightforward modifications of those for the causal systems, with the usual split between calculi with and without (*OR*).

For $\mathcal{SC}_{\mathbf{OUT}_2}$ and $\mathcal{SC}_{\mathbf{OUT}_4^+}$ the input and output sequents in the separation principles reflect the transition rules (*IN*) and (*OUT*). In $\mathcal{SC}_{\mathbf{OUT}_2^+}$ and $\mathcal{SC}_{\mathbf{OUT}_4^+}$, however, there is a single transition rule (*THROU*), which handles both input and output, so the two sequents merge into a single sequent (which we will call output sequent for convenience).

For $\mathcal{SC}_{\mathbf{OUT}_4^+}$, the outputs $\{X_j\}_{j \in J}$ will appear twice, but contraction in **LK** removes this redundancy. Thus $\mathcal{SC}_{\mathbf{OUT}_2^+}$ and $\mathcal{SC}_{\mathbf{OUT}_4^+}$ satisfy the same separation principle, recovering the equivalence of the logics $\mathbf{OUT}_2^+$ and $\mathbf{OUT}_4^+$ with classical logic noted in [MvdT01].

Lemma 32 (Separation principles for $\mathcal{SC}_{\mathbf{OUT}_2^+}$ and $\mathcal{SC}_{\mathbf{OUT}_4^+}$). *For both $k \in \{2, 4\}$, $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\mathbf{OUT}_k^+}$ iff $B, \{X_j\}_{j \in J} \vdash Y, \{A_i\}_{i \in I}$ (output sequent) is derivable in **LK** for every partition $(I, J) \in \mathcal{P}(\{1, \dots, n\})$.*

Proof. Analogous to the proofs of Lem. 17 and Lem. 19. The only difference is that the final sequent can be concluded only by (*THROU*); the derivability of the goal sequent is equivalent to the output sequent in the lemma (for both $\mathcal{SC}_{\mathbf{OUT}_2^+}$ and $\mathcal{SC}_{\mathbf{OUT}_4^+}$). $\square$

Notice that this characterization produces exactly the same 2^n sequents that arise in **LK** by applying bottom up the right implication rule and n times the left implication rule to $\{A_i \rightarrow X_i\}_{i=1}^n \vdash (B \rightarrow Y)$. By the invertibility of the **LK** rules, derivability of these 2^n sequents is equivalent to the derivability of $\{A_i \rightarrow X_i\}_{i=1}^n \vdash (B \rightarrow Y)$. The separation principles immediately yield the result from [MvdT01] that both $\mathbf{OUT}_2^+$ and $\mathbf{OUT}_4^+$ collapse into classical logic.

Corollary 2. *For both $k \in \{1, 3\}$, $\{(A_i, X_i)\}_{i=1}^n \vdash_{\mathbf{OUT}_k^+} (B, Y)$ iff $\{A_i \rightarrow X_i\}_{i=1}^n$ classically entails $(B \rightarrow Y)$.*

Hence, admissibility of the structural rules reduces to their admissibility in **LK**.

For $\mathcal{SC}_{\mathbf{OUT}_1^+}$ and $\mathcal{SC}_{\mathbf{OUT}_3^+}$ the transition rules correspond to the ex-falso and output sequents. Thus, in the separation principles for $\mathcal{SC}_{\mathbf{OUT}_1^+}$ and $\mathcal{SC}_{\mathbf{OUT}_3^+}$ these two sequents merge into a single sequent, which we will call the output sequent for uniform treatment later.

Lemma 33 (Separation principles for $\mathcal{SC}_{\mathbf{OUT}_1^+}$ and $\mathcal{SC}_{\mathbf{OUT}_3^+}$). *The following two equivalences hold.*

- $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\mathbf{OUT}_1^+}$ iff for every partition $(I, J) \in \mathcal{P}(\{1, \dots, n\})$, either $B, \{X_j\}_{j \in J} \vdash Y$ (output sequent) or $B \vdash A_i$ (input sequent) for some $i \in I$ is derivable in **LK**.

- $(A_1, X_1), \dots, (A_n, X_n) \vdash (B, Y)$ is derivable in $SC_{\mathbf{OUT}_3^+}$ iff for every partition $(I, J) \in \mathcal{P}(\{1, \dots, n\})$, either $B, \{X_j\}_{j \in J} \vdash Y$ (output sequent) or $B, \{X_j\}_{j \in J} \vdash A_i$ (input sequent) for some $i \in I$ is derivable in **LK**.

Proof. Analogous to the proofs of Lem. 22 and Lem. 23. The only difference is that the calculus now ends with the single transition rule (THROU) instead of both (IN) and (OUT). Consequently, the separation principle for $SC_{\mathbf{OUT}_1^+}$ and $SC_{\mathbf{OUT}_3^+}$ contains one merged output sequent rather than separate output and ex-falso sequents. $\square$

Using the separation principles, admissibility of the structural rules in $SC_{\mathbf{OUT}_1^+}$ and $SC_{\mathbf{OUT}_3^+}$ follows by a straightforward adaptation of the proofs for their causal counterparts.

Lemma 34. *The structural rules (IO-Wk), (IO-Ctr), and (IO-Cut) are admissible in $SC_{\mathbf{OUT}_1^+}$ and $SC_{\mathbf{OUT}_3^+}$.*

Proof. Analogous to the proof of Lem. 24, following the same three-step cut-elimination strategy from Lem. 18. The only difference is the absence of the ex-falso sequent: it is absorbed into the output sequent, which requires one additional cut in the final step. To illustrate, consider $SC_{\mathbf{OUT}_1^+}$ and compare with the proof for $SC_{\mathbf{OUT}_1^\perp}$ in Lem. 24. Since the ex-falso sequent (*ExFalsoS*) is removed, the target output sequent (*OutS*) is $B, \{X_j\}_{j \in J''}, \{W_j\}_{j \in J'} \vdash Y$. **Step 1** proceeds exactly as before (minus the ex-falso case), ending with (1) $B \vdash A_0$. **Step 2** proceeds exactly as before (minus the ex-falso case), ending with the derivable input sequent (2) $A_0, \{W_j\}_{j \in J'} \vdash X_0$. Finally, on **Step 3**, the input-sequent cases are again handled by weakening in **LK**. The remaining output case $B, X_0, \{X_j\}_{j \in J''} \vdash Y$ is now handled by two cuts in **LK**, first with (2) and then with (1), followed by a contraction of B , yielding (*OutS*). $\square$

Soundness and completeness of the calculi for the throughput I/O logics follow as in the causal case. The only additional steps are to show derivability of the new I/O principle (ID) and the new transition rule (THROU).

Theorem 31 (Soundness and completeness of $SC_{\mathbf{OUT}_k^+}$). *For $k \in \{1, \dots, 4\}$, $G \vdash (B, Y)$ is derivable in $SC_{\mathbf{OUT}_k^+}$ iff $G \vdash_{\mathbf{OUT}_k^+} (B, Y)$.*

Proof. Soundness. For each $k \in \{1, \dots, 4\}$, the rule (E_k) also appears in $SC_{\mathbf{OUT}_k}$, where its derivability from the I/O principles of $\mathbf{OUT}_i$ (which are included in $\mathbf{OUT}_i^+$) has already been established. For the rule (THROU), assuming that $Y \in Cn(\{B\})$, we derive (B, Y) by applying (WO) to the pair (B, B) obtained via (ID).

Completeness. Any derivation in $SC_{\mathbf{OUT}_i^\perp}$ can be simulated in $SC_{\mathbf{OUT}_i^+}$ by replacing applications of the rules (IN) and (OUT) with the stronger rule (THROU). Therefore, completeness of the calculi reduces to showing derivability of the principle (ID), which follows since the pair (A, A) can be obtained by a single application of (THROU). $\square$

5.5 Bridge: Reducing deontic I/O entailment to causal I/O entailment

Since the original deontic I/O logics and their causal counterpart differ only by rule (Bot), the entailment problem in the causal I/O logic can be reduced to the entailment problem in the corresponding original logic simply by adding $(\perp, \perp)$ to the set of the deriving pairs. By comparing the separation principles for these two families of logics, we obtain a reverse reduction. It shows that entailment in the original logic can be obtained from entailment in the causal logic by placing the additional condition of classical entailment of outputs.

Theorem 32. *For any $k \in \{1, \dots, 4\}$, $(A_1, X_1), \dots, (A_n, X_n) \vdash_{\mathbf{OUT}_k} (B, Y)$ iff both $(A_1, X_1), \dots, (A_n, X_n) \vdash_{\mathbf{OUT}_k^\perp} (B, Y)$ and $Y \in \text{Cn}(\{X_1, \dots, X_n\})$.*

Proof. The argument follows directly from the separation principles for the respective logics, summarized in Tab. 5.2.

(Left-to-right). Since each causal system extends its corresponding original I/O logic, the entailment already holds for $\mathbf{OUT}_k^\perp$. Moreover, instantiating the separation principle for $\mathbf{OUT}_k$ with the partition $(\emptyset, \{1, \dots, n\})$ implies the derivability of $X_1, \dots, X_n \vdash Y$ in **LK**.

(Right-to-left). $(A_1, X_1), \dots, (A_n, X_n) \vdash_{\mathbf{OUT}_k} (B, Y)$ follows from the separation principle. Every partition (I, J) with $I \neq \emptyset$ for $\mathbf{OUT}_i$ is handled by the same case of the separation principle for $\mathbf{OUT}_i^\perp$ (notice that for logics $\mathbf{OUT}_1^\perp/\mathbf{OUT}_3^\perp$, if the ex-falso sequent is derivable then one of the input sequents is also derivable). For the partition $(\emptyset, \{1, \dots, n\})$, the separation principle requires derivability of the output sequent $X_1, \dots, X_n \vdash Y$, which is given by the bridging condition. $\square$

Remark 8. This “bridging” condition $Y \in \text{Cn}(\{X_1, \dots, X_n\})$ appears already in embeddings of $\mathbf{OUT}_2$ and $\mathbf{OUT}_4$ into classical and normal modal logics in [MvdT00]. However, it appears that this simple condition was never explicitly proposed as the general syntactic link between original I/O logics and their causal counterparts.

5.6 Semantics of Input/Output models

The contrapositive reading of the separation principles yields a possible-world semantics for all I/O logics. It generalizes Bochman’s bimodel semantics [Boc04] and offers a simple, and modular semantic framework usable for exploring connections between I/O logics and other logical formalisms. Moreover, it subsumes existing semantic accounts for the original, causal, and throughout variants of I/O logics.

To introduce this semantics, we begin with the simplest case: the causal logic $\mathbf{OUT}_2^\perp$. According to the separation principle for $\mathbf{OUT}_2^\perp$ (Tab. 5.2), $\{(A_i, X_i)\}_{i=1}^n \not\vdash_{\mathbf{OUT}_2^\perp} (B, Y)$ iff there exists a partition $(I, J) \in \mathcal{P}(\{1, \dots, n\})$ such that both $B \vdash \{A_i\}_{i \in I}$ and

I/O inference principle	Frame property
(BOT)	$ In \geq 1$
(OR)	$ In \leq 1$
(CT)	$In \subseteq Out$
(ID)	$In \supseteq Out$

Figure 5.2: Frame conditions of I/O models corresponding to I/O inference principles.

$\{X_j\}_{j \in J} \vdash Y$ are not invalid. This means that there exist Boolean assignments ρ_1 and ρ_2 , with ρ_1 satisfying B but falsifying all formulas $\{A_i\}_{i \in I}$, and ρ_2 satisfying all formulas $\{X_j\}_{j \in J}$ but falsifying Y . Inspired by the bimodel semantics, we can treat this pair of Boolean assignments as a model representing an input state (ρ_1) and an output state (ρ_2), and say that an I/O pair (A, X) is satisfied by the model when the satisfaction of A in the input state implies the satisfaction of X in the output state. Notice that the separation principle then states exactly that (ρ_1, ρ_2) is a countermodel for the entailment under consideration. Indeed, for every pair (A_i, X_i) either $\rho_1 \not\models A_i$ or $\rho_2 \models X_i$, and hence $(\rho_1, \rho_2) \models (A_i, X_i)$. On the other hand $\rho_1 \models B$ and $\rho_2 \not\models Y$, so $(\rho_1, \rho_2) \not\models (B, Y)$. Therefore (ρ_1, ρ_2) satisfies all premises of the entailment but not the conclusion. Hence this notion of model is adequate for $OUT_2^\perp$.

Models for the remaining I/O logics can be reconstructed from their separation principles in the same way. In general, however, separation principles may involve several input sequents, which requires considering multiple input states. For this reason, in the general definition we allow both input and output states to be sets of worlds. We call the resulting structures *I/O models*.

Definition 30. A I/O model is a pair (In, Out) of sets of Boolean assignments, called input and output worlds. A set of Boolean assignments $\mathcal{S}$ validates a formula A (denoted $\mathcal{S} \models A$) when all assignments in $\mathcal{S}$ satisfy A . An I/O model (In, Out) satisfies a pair (A, X) (denoted as $(In, Out) \models (A, X)$) when either $In \not\models A$ or $Out \models X$.

I/O models yield semantic characterizations of all the I/O logics considered here, via simple frame conditions corresponding to each inference principle as show in Tab. 5.2.

Theorem 33. For any I/O logic $\mathcal{L}$, $G \vdash_{\mathcal{L}} (B, Y)$ iff for any I/O model M satisfying the corresponding frame properties for $\mathcal{L}$ from Tab. 5.2, if $M \models C$ for all $C \in G$ then $M \models (B, Y)$.

Proof. (Soundness). It is easy to check that every basic inference principle of I/O logics ((TOP), (WO), (SI), (AND)) is valid in all I/O models, and every additional inference principle ((BOT), (OR), (CT), (ID)) is valid in all I/O models satisfying the corresponding frame condition. Let us consider (CT) as an example. For any I/O model $M = (In, Out)$ with $In \subseteq Out$, suppose $M \models (A, X)$ and $M \models (A \wedge X, Y)$, then if $In \models A$ then $Out \models X$, then $In \models A \wedge X$ (since $In \subseteq Out$), then $Out \models Y$, implying that $M \models (A, Y)$.

(Completeness) Suppose towards contradiction that $G \not\models_{\mathcal{L}} (B, Y)$. By the separation principle for $\mathcal{L}$ (Tab. 5.2) there exists a partition (I, J) and boolean assignments falsifying the corresponding input sequents, the ex-falso sequent (if present), and the output sequent. We use these assignments to construct an I/O model M satisfying all conditionals in G , but falsifying (B, Y) . The model M is constructed as follows:

- Assignments falsifying the input sequents and the ex-falso sequent (if any) are taken as input worlds,
- Assignment falsifying the output sequent is taken as the output world.

If $\mathcal{L}$ includes (CT), we additionally include all input worlds among the output worlds. If $\mathcal{L}$ includes (ID), we additionally include all output worlds among the input worlds. Let us show that M satisfies all conditionals in G . Let $(A_k, X_k) \in G$. If $k \in I$ then there is an input sequent (or, for $\mathbf{OUT}_2^+$ and $\mathbf{OUT}_4^+$, an output sequent) with A_k on the RHS, which means that some input world in M falsifies A_k , so $M \models (A_k, X_k)$. If $k \in J$ then X_k occurs on the LHS of the output sequent (and, for logics with (CT), also on the LHS of the input and of the ex-falso sequents); consequently every output world in M satisfies X_k , and again $M \models (A_k, X_k)$. Now let us show that M falsifies (B, Y) . B occurs on the LHS of all sequents corresponding to input worlds, while Y occurs on the RHS of the output sequent. Thus all input worlds satisfy B , while some output world falsifies Y , therefore $M \not\models (B, Y)$. Finally, let us show that M satisfies the frame conditions associated with $\mathcal{L}$ (Tab. 5.2). Conditions for (CT) and (ID) are ensured directly by the choice of worlds, for logics with (BOT) there is always at least one sequent corresponding to input worlds (the ex-falso sequent for $\mathbf{OUT}_1^\perp/\mathbf{OUT}_3^\perp$, the only input sequent for $\mathbf{OUT}_2^\perp/\mathbf{OUT}_4^\perp$, and the output sequent for $\mathbf{OUT}_1^+/\mathbf{OUT}_2^+/\mathbf{OUT}_3^+/\mathbf{OUT}_4^+$), and for logics with (OR) there is always at most one sequent corresponding to input worlds (the only input sequent for $\mathbf{OUT}_2^\perp/\mathbf{OUT}_2^\perp/\mathbf{OUT}_4^\perp/\mathbf{OUT}_4^\perp$, and the output sequent for $\mathbf{OUT}_2^+/\mathbf{OUT}_4^+$). $\square$

Remark 9. *From the construction in the proof above follows that the size of the I/O models can be bounded: $|In| \leq 1$ for logics with (OR) and $|In| \leq |G| + 2$ otherwise, $|Out| = 1$ for logics without (CT) and $|Out| \leq |G| + 2$ otherwise.*

Let us see how our semantics can be interpreted in the normative context.

Example 13. *Consider the normative theory from Example 11:*

$$\{(\top, \text{Lawful}), (\neg \text{Lawful}, \text{Erase}), (\text{Lawful}, \neg \text{Erase})\}.$$

We show that the unethical obligation $(\neg \text{Lawful}, \text{KillMary})$ cannot be derived in $\mathbf{OUT}_1$ and $\mathbf{OUT}_2$. An I/O countermodel (In, Out) to this entailment should be such that (a) all input worlds satisfy $\neg \text{Lawful}$, (b) some output world does not satisfy KillMary and (c) for every conditional obligation (A, X) in the normative theory either all output worlds satisfy X or there is an input world that does not satisfy A . Our negative answer is based on the countermodel $(\{in\}, \{out\})$, such that boolean valuations in and out satisfy

$\neg \text{Lawful}$ and $\{\neg \text{KillMary}, \text{Lawful}, \text{Erase}\}$, respectively. Intuitively *out* is an “ideal” world as it satisfies all conditional obligations triggered in the given situation (and in which *KillMary* does not happen), while *in* describes a case consistent with the given situation which explains why $(\text{Lawful}, \neg \text{Erase})$ is not triggered.

5.7 Bridge: I/O models and existing semantical descriptions for I/O logics

The new semantics of I/O models can be related to the existing semantic frameworks proposed for the different logics in the I/O family: bimodel semantics for the causal I/O logics [Boc04] and the operational semantics for the original I/O logics [MvdT00].²

I/O models vs bimodels. Bimodel semantics was introduced in [Boc04] for the general case of $\text{OUT}_1^\perp$ as pairs (α, β) of consistent, deductively closed sets of propositional formulas (i.e., sets containing all their classical consequences and not containing $\perp$). An I/O pair (A, B) is satisfied in such a model if $A \in \alpha$ implies $B \in \beta$. As noted already in [Boc04], this definition admits a possible-worlds reading: replacing each deductively closed set with the set of worlds satisfying all formulas in them transforms a bimodel into an I/O model (Def. 30) that satisfies the same I/O pairs. Conversely, any I/O model with non-empty *In* and *Out* can be converted into a bimodel by replacing each world set with the set of formulas satisfied by all its worlds. Under this correspondence, the frame condition for the principle (*BOT*) matches the consistency requirement for the input set α , while the conditions for additional principles (*CT*) and (*OR*) from [Boc04] directly correspond to the I/O model frame conditions in Tab. 5.2. Specifically, for $\text{OUT}_2^\perp$, both α and β must be complete theories, corresponding to singleton input and output worlds ($|In| = |Out| = 1$), since one output world suffices in the absence of (*CT*), while $\text{OUT}_3^\perp$ is characterized by *inclusive* bimodels, defined by $\beta \subseteq \alpha$, which corresponds to $In \subseteq Out$. The soundness and completeness of bimodel semantics with respect to the I/O principles were established in [Boc04] through canonical model construction. The I/O models introduced in Def. 30 can be seen as a more constructive perspective on bimodels, independently derived from the separation principles and generalized to all core I/O logics.

I/O models vs procedural semantics. The original I/O logics were conceived as a means of “logically assisted transformation” [MvdT00], and accordingly they were equipped with a *procedural* semantic account: for each original I/O logic $\mathcal{L}$, they defined an operator (abusively) denoted by $\mathcal{L}$, which takes a set of conditionals G and a set of input formulas α , and returns a set of output formulas $\mathcal{L}(G, \alpha)$. These operators for the

²A formal connection between these two existing semantics was previously unknown. As noted in [vdTP13]: “As a matter of fact, there is no direct (formal) connection between Bochman’s semantics and the operational semantics for I/O logic. The linkage is established only via the axiomatic characterization: both semantics yield essentially the same axiom system”.

original I/O logics were defined as follows (here $G(\alpha) = \{B \mid (A, B) \in G, A \in \alpha\}$ and a *complete* set is either a subset-maximal set among classically consistent sets or the set of all Boolean formulas):

$$\begin{aligned} \mathbf{OUT}_1(G, \alpha) &= \cap \{Cn(G(\beta)) \mid \alpha \subseteq \beta = Cn(\beta)\} \\ \mathbf{OUT}_2(G, \alpha) &= \cap \{Cn(G(\beta)) \mid \alpha \subseteq \beta = Cn(\beta), \beta \text{ is complete}\} \\ \mathbf{OUT}_3(G, \alpha) &= \cap \{Cn(G(\beta)) \mid \alpha \subseteq \beta = Cn(\beta) \supseteq G(\beta)\} \\ \mathbf{OUT}_4(G, \alpha) &= \cap \{Cn(G(\beta)) \mid \alpha \subseteq \beta = Cn(\beta) \supseteq G(\beta), \beta \text{ is complete}\} \end{aligned}$$

Operators for the throughput logics can be derived from their original versions using the following equation [MvdT01]: $\mathbf{OUT}_k^+(G, \alpha) = \mathbf{OUT}_k(G \cup I, \alpha)$, where I is the set of pairs (A, A) for each propositional formula A . This semantic characterization was proven in [MvdT00] to be sound and complete for finite input sets with respect to the syntactical derivation using I/O principles, establishing that $G \vdash_{\mathcal{L}} (B, Y)$ if and only if $Y \in \mathcal{L}(G, \{B\})$.

There is a direct correspondence between the negative reading of these definitions and I/O countermodels. Namely, $Y \in \mathcal{L}(G, \{B\})$ is equivalent to the existence of set β that satisfies the conditions from the definition of $\mathcal{L}$ above (including $\{B\} \subseteq \beta$) such that $Y \notin Cn(G(\beta))$. From such β we can reconstruct the following countermodel for entailment of (B, Y) from G : take (In, Out) where In (resp. Out) contains all Boolean assignments satisfying all formulas in β (resp. in $G(\beta)$), and notice $(In, Out) \not\models (B, Y)$ and $(In, Out) \models (A, X)$ for every $(A, X) \in G$ (since either $A \notin \beta$ or $X \in G(\beta)$). Conversely, we can reconstruct suitable β from any countermodel (In, Out) by taking all formulas A such that $In \models A$. Under this correspondence, the conditions on β transform exactly to frame conditions of I/O models from Tab. 5.2: completeness of β is equivalent to the fact it is satisfied by at most one assignment, $Cn(\beta) \supseteq G(\beta)$ is equivalent to the fact $In \subseteq Out$. Also, the same correspondence will work if we replace G with $G \cup I$ for throughput logics, yielding models with the required frame condition $In \supseteq Out$.

5.8 Bridge: Embeddings of I/O logics into classical and normal modal logics

In this section, we use the I/O models to construct embeddings of I/O logics into classical propositional logic and normal modal logics. These embeddings clarify the position of I/O logics within well-established logical frameworks and enable the direct application of existing reasoning techniques for the target logics; in particular, the classical embedding allows the use of SAT solvers for automated deduction for I/O logics.

The original paper [MvdT00] already presented embeddings of the logics $\mathbf{OUT}_2$ and $\mathbf{OUT}_4$ into both classical propositional logic and normal modal logic (specifically, $\mathbf{K}$ for $\mathbf{OUT}_2$ and $\mathbf{T}$ for $\mathbf{OUT}_4$). However, these constructions crucially rely on the presence of complete sets in the procedural semantics, and thus apply only to I/O logics that validate

the (Or) principle. This reliance even led to the conjecture³ that no straightforward embedding could be achieved for the I/O logics without (OR). We show that I/O models and their frame conditions admit natural and uniform encodings into both classical propositional and normal modal logics.

Embedding into classical logic

Our embedding naturally extends (and slightly modifies) the embeddings in [MvdT00]. We start by encoding the notion of I/O countermodel within classical propositional logic. To this end, we introduce a distinct copy of each propositional variable to represent its value in different worlds of the model. Formally, we define a family of injective mappings on propositional atoms $\{\chi_i: Atm \rightarrow Atm\}_{i \in \mathbb{N}_0}$ such that $Dom(\chi_i) \cap Dom(\chi_j) = \emptyset$ for all $i \neq j$ (such a family exists as the set Atm is infinite). We encode the value of a variable x in world i using $\chi_i(x)$, with world 0 designated as the output world that falsifies the non-entailed conditional. The satisfaction of a formula A in world i is then represented by a copy of A in which each variable is replaced by its image under χ_i ; we denote this transformed formula by $\chi_i(A)$. The satisfaction of a pair (A, X) can be encoded as $(\bigwedge_{i \in In} \chi_i(A)) \rightarrow (\bigwedge_{i \in Out} \chi_i(X))$. By varying the number of input worlds and/or adding input indices to output indices and vice versa, we can capture the logics **OUT**₁[⊥]–**OUT**₄[⊥] and **OUT**₁⁺–**OUT**₄⁺. Modifying this embedding to represent the original logics is less straightforward. However, we can use our reduction of entailment in the original I/O logics to entailment in the causal I/O logics (Th. 32), and include the “bridging condition” from there as a side condition in the embedding. The resulting embeddings into classical logics for all considered I/O logics are presented in Tab. 5.3.

Let us denote by $cl\text{-}emb(\mathcal{L}, n, (A, X))$ the embedding of the pair (A, X) from the I/O logic $\mathcal{L}$ as in Tab. 5.3.

Theorem 34. *Let $\mathcal{L}$ be a core I/O logic. $(A_1, X_1), \dots, (A_n, X_n) \vdash_{\mathcal{L}} (B, Y)$ iff*

- (a) $cl\text{-}emb(\mathcal{L}, n, (B, Y)) \in Cn(\{cl\text{-}emb(\mathcal{L}, n, (A_i, X_i))\}_{i=1}^n)$, when $\mathcal{L}$ is a causal or a throughput I/O logic,
- (b) $cl\text{-}emb(\mathcal{L}, n, (B, Y)) \in Cn(\{cl\text{-}emb(\mathcal{L}, n, (A_i, X_i))\}_{i=1}^n)$ and $Y \in Cn(\{X_i\}_{i=1}^n)$, when $\mathcal{L}$ is an original I/O logic.

Proof. (b) follows from (a) by Th. 32, so we only need to prove (a). We argue contrapositively.

(Right-to-left) Suppose towards contradiction that $\{cl\text{-}emb(\mathcal{L}, n, (A_n, X_n))\}_{i=1}^n$ does not entail $cl\text{-}emb(\mathcal{L}, n, (B, Y))$ in classical logic. Then there is a boolean valuation ρ satisfying all formulas in $\{cl\text{-}emb(\mathcal{L}, n, (A_n, X_n))\}_{i=1}^n$ and falsifying $cl\text{-}emb(\mathcal{L}, n, (B, Y))$. Let us

³From [MvdT00]: “As far as the authors are aware, it is not possible to characterise the system of simple-minded output (with or without reusability) by relabeling or modal logic in a straightforward way. The (OR) rule appears to be needed, so that we can work with complete sets.”

I/O logic $\mathcal{L}$	Embedding $cl-emb(\mathcal{L}, n, (A, X))$	Bridging side condition
OUT₁	$(\bigwedge_{i=1}^n \chi_i(A)) \rightarrow \chi_0(X)$	yes
OUT₂	$\chi_1(A) \rightarrow \chi_0(X)$	yes
OUT₃	$(\bigwedge_{i=1}^n \chi_i(A)) \rightarrow (\bigwedge_{i=0}^n \chi_i(X))$	yes
OUT₄	$\chi_1(A) \rightarrow (\chi_0(X) \wedge \chi_1(X))$	yes
OUT₁[⊥]	$(\bigwedge_{i=1}^n \chi_i(A)) \rightarrow \chi_0(X)$	no
OUT₂[⊥]	$\chi_1(A) \rightarrow \chi_0(X)$	no
OUT₃[⊥]	$(\bigwedge_{i=1}^n \chi_i(A)) \rightarrow (\bigwedge_{i=0}^n \chi_i(X))$	no
OUT₄[⊥]	$\chi_1(A) \rightarrow (\chi_0(X) \wedge \chi_1(X))$	no
OUT₁⁺	$(\bigwedge_{i=0}^n \chi_i(A)) \rightarrow \chi_0(X)$	no
OUT₂⁺	$\chi_0(A) \rightarrow \chi_0(X)$	no
OUT₃⁺	$(\bigwedge_{i=0}^n \chi_i(A)) \rightarrow (\bigwedge_{i=0}^n \chi_i(X))$	no
OUT₄⁺	$\chi_0(A) \rightarrow \chi_0(X)$	no

Table 5.3: Embeddings of $(A_1, X_1), \dots, (A_n, X_n) \vdash_{\mathcal{L}} (B, Y)$ into classical propositional logic for every core I/O logic $\mathcal{L}$. The embeddings for **OUT₁[⊥]**-**OUT₄[⊥]** and **OUT₁⁺**-**OUT₄⁺** have the form $cl-emb(\mathcal{L}, n, (B, Y)) \in Cn(\{cl-emb(\mathcal{L}, n, (A_i, X_i))\}_{i=1}^n)$, where the function $cl-emb(\mathcal{L}, n, (A, X))$ maps an arbitrary pair (A, X) into a classical implication, and n is the number of deriving pairs. The embeddings for **OUT₁**-**OUT₄** require the additional “bridging” side condition $Y \in Cn(\{X_1, \dots, X_n\})$ (Th. 32).

consider the “projecting” boolean valuations ρ_i for each i defined by $\rho_i(x) = \rho(\chi_i(x))$. For $\mathcal{L} = \mathbf{OUT}_1^\perp$, consider the I/O model $M = (\{\rho_i\}_{i=1}^n, \rho_0)$. It is easy to see that $M \models (A, X)$ iff $\rho \models cl-emb(\mathbf{OUT}_1^\perp, n, (A, X))$, so M is a countermodel for entailment of (B, Y) from G . The same construction can be naturally adapted to all extensions of **OUT₁[⊥]**: for extensions with *(OR)* we replace $\{\rho_i\}_{i=1}^n$ with just $\{\rho_1\}$, for extensions with *(CT)* we add all input worlds to the set of output worlds; and for extensions with *(ID)* we add all output worlds to the set of input worlds.

(Left-to-right) Suppose towards contradiction that $\{(A_i, X_i)\}_{i=1}^n$ does not entail (B, Y) in $\mathcal{L}$, then there is an I/O model (In, Out) such that $M \not\models (B, Y)$ and $M \models (A_i, X_i)$ for all $i \in \{1, \dots, n\}$. Let us construct the boolean valuation ρ satisfying formulas $\{cl-emb(\mathcal{L}, n, (A_i, X_i))\}_{i=1}^n$ and falsifying $cl-emb(\mathcal{L}, n, (B, Y))$. Since $(In, Out) \not\models (B, Y)$ there is $\rho_0 \in Out$ such that $\rho_0 \not\models Y$; define ρ on $Dom(\chi_0)$ by $\rho(\chi_0(x)) = \rho_0(x)$. Now, for each $i \in \{1, \dots, n\}$, take any $in \in In$ such that $in \not\models A_i$ if such input worlds exist, otherwise take any $in \in In$ (there is at least one since we consider logics with *(BOT)*), and define ρ on $Dom(\chi_i)$ by $\rho(\chi_i(x)) = in(x)$. On variables outside $\bigcup_{i=0}^n Dom(\chi_i)$ we can define ρ arbitrarily. Thus defined ρ falsifies $cl-emb(\mathcal{L}, n, (B, Y))$ and satisfies $cl-emb(\mathcal{L}, n, (A_i, X_i))$ for each $i \in \{1, \dots, n\}$: ρ falsifies $\chi_0(Y)$ by choice of values on $Dom(\chi_0)$, ρ satisfies $\chi_i(B)$ for $i \in \{1, \dots, n\}$ (and also $i = 0$ for logics with *(ID)*) since the definition on $\bigcup_{i=1}^n Dom(\chi_i)$ is based on input worlds, and also for each $i \in \{1, \dots, n\}$, either $In \not\models A_i$ and then ρ falsifies $\chi_i(A_i)$ (and also $\chi_1(A_i)$ if $\mathcal{L}$ contains *(OR)* and $\chi_0(A_i)$ if $\mathcal{L}$ contains both *(OR)* and *(ID)*) or $Out \models X_i$ and then ρ satisfies $\chi_0(X_i)$ (and also

$\chi_j(X_i)$ for every $j \in \{1, \dots, n\}$ in case $\mathcal{L}$ contains (CT) . $\square$

Our embeddings for $\mathbf{OUT}_2^+$ and $\mathbf{OUT}_4^+$ reflect that in these logics I/O pairs collapse to classical implications, as observed in [MvdT01].

Example 14. *We encode the I/O entailment for $\mathbf{OUT}_3$*

$$(\top, Lawful), (\neg Lawful, Erase), (Lawful, \neg Erase) \vdash (\neg Lawful, KillMary)$$

from Example 11 in classical logic. The encoding uses propositional variables from the set $\bigcup_{i=0}^3 \{\chi_i(Lawful), \chi_i(Erase), \chi_i(KillMary)\}$ together with the original variables $\{Lawful, Erase, KillMary\}$, and consists of the two entailments that hold in classical logic:

(I) the mapping of the sequent: $Q \in Cn(N_1, N_2, N_3)$, where

$$\begin{aligned} Q &= (\bigwedge_{i=1}^3 \neg \chi_i(Lawful)) \rightarrow \bigwedge_{i=0}^3 \chi_i(KillMary), \\ N_1 &= (\bigwedge_{i=1}^3 \top) \rightarrow \bigwedge_{i=0}^3 \chi_i(Lawful), \\ N_2 &= (\bigwedge_{i=1}^3 \neg \chi_i(Lawful)) \rightarrow \bigwedge_{i=0}^3 \chi_i(Erase), \\ N_3 &= (\bigwedge_{i=1}^3 \chi_i(Lawful)) \rightarrow \bigwedge_{i=0}^3 \neg \chi_i(Erase), \end{aligned}$$

(II) the bridging side condition: $KillMary \in Cn(\{Lawful, Erase, \neg Erase\})$.

Remark 10. *The embedding for $\mathbf{OUT}_2$ (including the bridging condition), known as “relabeling”, was originally presented in [MvdT00]. For $\mathbf{OUT}_4$, [MvdT00] provides a different embedding that extends the one for $\mathbf{OUT}_2$ by adding classical implications to capture (CT) .*

Embeddings into modal logics

To construct the embeddings into modal logics, we establish a connection between I/O models and pointed relational models. Specifically, we associate a world w in a pointed relational model (M, w) with the output world, and interpret the set of all worlds accessible from w as the set of input worlds. Under this correspondence, the satisfaction of (A, X) in $(In, \{out\})$ is equivalent to the satisfaction of the modal formula $\Box A \rightarrow X$ in the pointed model (M, w) . Moreover, the frame conditions from Tab. 5.2 (apart from the one corresponding to the principle (CT)) align with well-known frame conditions in relational semantics:

- the condition $|In| \geq 1$ corresponds to seriality (at least one world is accessible from each world), characterized by the axiom $(D) : \Box A \rightarrow \Diamond A$;
- $|In| \leq 1$ corresponds to functionality (at most one world is accessible from each world), characterized by the axiom $(F) : \Diamond A \rightarrow \Box A$;

I/O logic $\mathcal{L}$	Modal embedding $\Box\text{-emb}(\mathcal{L}, (A, X))$	Target modal logic $\mathcal{L}^\Box$
OUT₁	$(\Box A \rightarrow X)$	K
OUT₂	$(\Box A \rightarrow X)$	KF
OUT₃	$(\Box A \rightarrow (X \wedge \Box X))$	K
OUT₄	$(\Box A \rightarrow (X \wedge \Box X))$	KF
OUT₁[↓]	$(\Box A \rightarrow X)$	KD
OUT₂[↓]	$(\Box A \rightarrow X)$	KDF
OUT₃[↓]	$(\Box A \rightarrow (X \wedge \Box X))$	KD
OUT₄[↓]	$(\Box A \rightarrow (X \wedge \Box X))$	KDF
OUT₁⁺	$(\Box A \rightarrow X)$	KT
OUT₂⁺	$(\Box A \rightarrow X)$	KTF
OUT₃⁺	$(\Box A \rightarrow (X \wedge \Box X))$	KT
OUT₄⁺	$(\Box A \rightarrow (X \wedge \Box X))$	KTF

Table 5.4: Embeddings of the core I/O logics into normal modal logics.

- $Out \subseteq In$ corresponds to reflexivity (each world is accessible from itself), characterized by the axiom $(T) : \Box A \rightarrow A$.

As it is unclear how to characterize the condition $In \subseteq Out$ by means of frame conditions on a corresponding relational model, to cover logics with (CT) we instead modify the embedding of pairs (A, B) to $(\Box A \rightarrow X \wedge \Box X)$. With this modification we can treat all accessible worlds as output worlds.

Tab. 5.4 summarizes the embeddings of all considered I/O logics $\mathcal{L}$ into the corresponding target modal logics $\mathcal{L}^\Box$. The latter turn out to be either **K** or one of its extensions⁴ with axiom (D) , axiom (F) , and/or axiom (T) , depending on the I/O logic under consideration.

Theorem 35. *For any core I/O logic $\mathcal{L}$, $G \vdash_{\mathcal{L}} (B, Y)$ iff $\{\Box\text{-emb}(\mathcal{L}, (A, X))\}_{(A, X) \in G}$ entails $\Box\text{-emb}(\mathcal{L}, (B, Y))$ in $\mathcal{L}^\Box$.*

Proof. We argue by contraposition in both directions, transforming countermodels between the two semantics.

(Left-to-right) Suppose towards contradiction that $\Box\text{-emb}(\mathcal{L}, (B, Y))$ is not entailed by $\{\Box\text{-emb}(\mathcal{L}, (A, X))\}_{(A, X) \in G}$ in $\mathcal{L}^\Box$. Then there exists a relational model $M = \langle W, R, V \rangle$ and a world $w_0 \in W$ such that $M, w_0 \not\models \Box\text{-emb}(\mathcal{L}, (B, Y))$ and $M, w_0 \models \Box\text{-emb}(\mathcal{L}, (A, X))$ for every $(A, X) \in G$. We transform (M, w_0) into an I/O model that falsifies the entailment. For logics without (CT) , we take the model $M' = (In, \{out\})$ where $In = \{V(w) \mid w \in W, w_0 R w\}$ and $out = V(w_0)$. By inspection of the embedding, for every pair (A, X) , $(In, \{out\}) \models (A, X)$ is equivalent to $M, w_0 \models (\Box A \rightarrow X)$ for any I/O pair (A, X) . Also, the frame properties of M imply the corresponding frame

⁴Each modal logic is named by appending the corresponding axiom letters to **K**.

properties of $(In, \{out\})$: if M satisfies functionality then $|In| \leq 1$, if M satisfies seriality then $|In| \geq 1$, and if M satisfies reflexivity then $out \in In$. Thus the constructed I/O model satisfies the semantic constraints of $\mathcal{L}$, and $(In, \{out\})$ falsifies the entailment $G \vdash_{\mathcal{L}} (B, Y)$ by Th. 33. For logics with (CT) , we take the I/O model $(In, \{out\} \cup In)$ with In and out defined as before. Now $(In, \{out\} \cup In) \models (A, X)$ is equivalent to $M, w_0 \models (\Box A \rightarrow X \wedge \Box X)$ for any I/O pair (A, X) , and the frame properties of M again imply the corresponding frame properties of $(In, \{out\} \cup In)$. Hence $(In, \{out\} \cup In)$ is a countermodel for the entailment of $G \vdash_{\mathcal{L}} (B, Y)$.

(Right-to-left) Assume towards contradiction that G does not entail (B, Y) in $\mathcal{L}$, then by Th. 33 there is an I/O model (In, Out) such that $(In, Out) \not\models (B, Y)$ and for every $(A, X) \in G$, $(In, Out) \models (A, X)$. We construct a relational model $\langle W, R, V \rangle$ as follows: $W = In \cup Out$, $V(\rho) = \rho$ for every $\rho \in W$, and $\rho R \rho'$ iff $\rho' \in In$ and either $\rho \in Out$ or $\rho = \rho'$. Notice that $(In, Out) \models (A, X)$ is equivalent to the fact that $(\Box A \rightarrow X)$ is satisfied in $\langle W, R, V \rangle$ in all worlds from Out , and for logics with (CT) it is also equivalent to the fact that $(\Box A \rightarrow X \wedge \Box X)$ is satisfied in all worlds from Out (since for such logics $In \subseteq Out$). Therefore, there is a world w in Out , such that $\langle W, R, V \rangle, w \not\models \Box\text{-emb}(\mathcal{L}, (B, Y))$ and $\langle W, R, V \rangle, w \models \Box\text{-emb}(\mathcal{L}, (A, X))$ for all $(A, X) \in G$. Also the frame properties of (In, Out) (apart from the property of (CT)) imply the corresponding frame properties of $\langle W, R, V \rangle$: if $|In| \geq 1$ then R satisfies seriality, if $|In| \leq 1$ then R satisfies functionality, and if $In \supseteq Out$ then R satisfies reflexivity. Thus, $\langle W, R, V \rangle$ is a countermodel for the entailment of $\Box\text{-emb}(\mathcal{L}, (B, Y))$ from $\{\Box\text{-emb}(\mathcal{L}, (A, X))\}_{(A, X) \in G}$ in $\mathcal{L}^\Box$. $\square$

Remark 11. The embeddings for **OUT₂** and **OUT₄** presented in [MvdT00] are different: a pair (A, X) is translated into a formula $(A \rightarrow \Box X)$ within the logic **K** (for **OUT₂**) and **KT** (for **OUT₄**). Additionally, the “bridging condition” (Th. 32) is required to hold, as in the embeddings into classical logic described above (so these translations, in effect, provide embeddings for the logics **OUT₂[⊥]** and **OUT₄[⊥]** rather than for **OUT₂** and **OUT₄**). Note that moving the $\Box$ modality from outputs to inputs yields a more fine-grained embedding that distinguishes between different I/O logics, and can also capture those without (OR) . Indeed, for sequents of the form $(A \rightarrow \Box X), \dots, (A \rightarrow \Box X) \models (B \rightarrow \Box Y)$ validity in **K**, **KF**, **KD** and **KDF** coincides, and in each case corresponds to entailment in **OUT₂[⊥]**. By contrast, for sequents of the form $(\Box A_1 \rightarrow X_1), \dots, (\Box A_n \rightarrow X_n) \models (\Box B \rightarrow Y)$ validity in **K** corresponds to **OUT₁**, in **KF** to **OUT₂**, in **KD** to **OUT₁[⊥]**, and in **KDF** to **OUT₂[⊥]**.

5.9 Bridge: Reducing entailment in Åqvist's logic **E** to I/O entailment

Our possible-world semantics for I/O logics makes it possible to connect them not only with normal modal logics, but also with preference-based deontic logics, and with Åqvist's systems in particular. This connection is based on the observation that, when considering entailment of conditional obligation $\bigcirc(B|A)$ from the set of conditional obligations $\{\bigcirc(B_i|A_i)\}_{i=1}^n$ in **E**, where A, B, A_i, B_i are propositional formulas, only a restricted

fragment of a countermodel is relevant: to falsify the entailment, it suffices to consider an A -best world violating B together with the worlds strictly preferred to it that witness why it is not best with respect to the antecedents A_i . These worlds already contain all information required to reconstruct a countermodel. This observation suggests a direct correspondence with I/O models: the falsifying world can be identified with the output world, while the strictly preferred worlds correspond to the input worlds.

However, conditional obligations in Åqvist's logics cannot simply be treated as pairs in I/O logics, since conditionals behave very differently in these two formalisms. In particular, I/O logics validate strengthening of input, whereas in Åqvist's systems it is crucial that the condition in a conditional obligation cannot generally be strengthened. To overcome this, we employ a more involved translation, which maps a conditional obligation $\bigcirc(A|B)$ into the I/O pair $(\neg A, \neg A \vee B)$. The intuition behind this translation is to assert with I/O pair that if all worlds strictly preferred to the considered world satisfy $\neg A$, then either the considered world must satisfy $\neg A$ as well, making it not an A -best world, or it must satisfy obligation B . Thus, whenever the world is A -best, the obligation B is enforced. This translation yields an embedding of conditional entailment in **E** into the logic **OUT₁**.

Theorem 36. *Let $A, B, \{A_i\}_{i=1}^n, \{B_i\}_{i=1}^n$ be propositional formulas. $\{\bigcirc(B_i|A_i)\}_{i=1}^n$ entails $\bigcirc(B|A)$ in Åqvist's logic **E** iff*

$$\{(\neg A_i, \neg A_i \vee B_i)\}_{i=1}^n \vdash_{\mathbf{OUT}_1} (\neg A, \neg A \vee B).$$

Proof. By countermodel transformation. Let M be a preferential model such that $M \models \bigcirc(B_i|A_i)$ for all $i \in \{1, \dots, n\}$ and $M \not\models \bigcirc(B|A)$, there should be a world u in M such that $u \in \max_{\succ}(|A|^M)$ and $M, u \not\models B$. It is easy to verify that I/O model with valuation of u as a single output world and valuations of all w such that $w \succ u$ in M satisfies pairs $(\neg A_i, \neg A_i \vee B_i)$ but not $(\neg A, \neg A \vee B)$. Conversely, given an I/O model (In, Out) , there is a valuation out_0 in Out such that $out_0 \not\models \neg A \vee B$. Consider the preferential model with the set of worlds consisting of a single world v_0 together with infinite sequence of worlds $\{u_i^{in}\}_{i=1}^\infty$ for each $in \in In$, with valuation $V(v) = out_0$ and $V(u_i^{in}) = in$ for each $in \in In$, and with the following preference relation: $x \succeq y$ when either $x = u_1^{in}$ and $y = v$ or $x = u_{i+1}^{in}$ and $y = u_i^{in}$. It is easy to verify that such model satisfies all obligations in $\{\bigcirc(B_i|A_i)\}_{i=1}^n$ (note that only world v can be A -best in such model construction) and falsifies $\bigcirc(B|A)$ in the world v . $\square$

Apart from the interaction of conditionals, another important feature of Åqvist's logics is the presence of alethic modalities $\Box D$, which can be used to impose global constraints on the model. We can incorporate alethic facts in our embedding too, translating them into I/O pairs of the form $(\top, D)$. To capture the global nature of such facts, however, we need to move from **OUT₁** to **OUT₃**, where outputs are required to hold throughout the entire model, not only in the output worlds. Note that the previous translation of conditional obligations is also valid in **OUT₃**, as entailment in **OUT₁** and **OUT₃** coincide for pairs of this form.

Theorem 37. *Let $A, B, \{A_i\}_{i=1}^n, \{B_i\}_{i=1}^n, \{D_i\}_{i=1}^m$ be propositional formulas. Set of formulas $\{\bigcirc(B_i|A_i)\}_{i=1}^n \cup \{\Box D_i\}_{i=1}^m$ entails $\bigcirc(B|A)$ in Åqvist's logic **E** iff*

$$\{(\neg A_i, \neg A_i \vee B_i)\}_{i=1}^n \cup \{(\top, D_i)\}_{i=1}^m \vdash_{\mathbf{OUT}_3} (\neg A, \neg A \vee B).$$

Proof. Using the same countermodel transformation as in the proof of Th. 36. Satisfaction of $\Box D$ in the preference model and of $(\top, D)$ in the I/O model equivalently means that D is satisfied in all world in the model. $\square$

This connection illustrates that, despite their comparatively simple and monotonic nature, I/O logics are capable of incorporating substantially richer forms of conditional reasoning. Through the above translation, they inherit reasoning patterns of preference-based deontic logics, including forms of contrary-to-duty reasoning.

5.10 Prover: IOentailment

A SAT-based theorem prover **IOENTAILMENT**⁵ for both original and causal I/O logics was developed by Marit Einböck in her Bachelor's thesis [Ein24]. The prover implements the embeddings of I/O entailment into classical propositional logic from Tab. 5.3 and supports the four original I/O logics **OUT**₁–**OUT**₄ and their causal counterparts **OUT**₁[⊥]–**OUT**₄[⊥]. Given a normative theory and a goal pair, the prover either confirms the entailment or returns an explicit I/O countermodel witnessing its failure.

At the implementation level, the input formulas are first parsed into syntax trees and subsequently transformed into directed acyclic graphs (DAGs), so that identical subformulas are represented only once and can be shared during the construction of the SAT encoding. The world-indexed copies of propositional variables used abstractly in Sec. 5.8 are implemented by systematic renaming; for instance, $x[w3]$ represents the value of x in world 3. The implementation is modular across the different I/O logics: common parts of the encoding are shared between the causal systems, while the treatment of outputs and the required number of worlds are specialized according to the selected logic. The original systems reuse the implementation of their causal counterparts together with the additional bridging condition given by Th. 32.

The resulting propositional formula is submitted to a SAT solver. Unsatisfiability establishes the requested I/O entailment, while a satisfying assignment witnesses non-entailment. In the latter case, **IOENTAILMENT** translates the assignment back into an I/O model: the original propositional variable and the world in which it is true can be recovered from the indexed variable names, with world 0 representing the output world. A web interface allows the user to enter the deriving and goal pairs, select one of the eight supported logics, and directly inspect either the positive entailment result or the generated countermodel.

⁵<https://github.com/marit-dino/IOentailmentWeb>

The implementation was evaluated in [Ein24] on two case studies from the literature, with the obtained results cross-checked against the corresponding sequent calculi. The first case study concerns a formalization of a fragment of the General Data Protection Regulation from [BPvdT20], discussed partially in our Ex. 11, involving obligations concerning lawful processing and erasure of personal data, and asks whether an unrelated “unethical” obligation can be derived. The example distinguishes the weaker systems $\mathbf{OUT}_1, \mathbf{OUT}_2, \mathbf{OUT}_1^\perp, \mathbf{OUT}_2^\perp$, in which the obligation is not entailed, from the reusable systems $\mathbf{OUT}_3, \mathbf{OUT}_4, \mathbf{OUT}_3^\perp, \mathbf{OUT}_4^\perp$, in which it is derivable.

The second case study, adapted from a formalization of traffic rules [CPS21], combines obligations to turn on the lights when it is dark or foggy, to drive when the lights are on, and not to drive on an icy road when it is dark or foggy. The prover checks whether the conflicting obligation to drive under the latter circumstances follows. It turns out to be entailed in the logics $\mathbf{OUT}_4$ and $\mathbf{OUT}_4^\perp$, and not entailed in the weaker I/O logics. In all tested cases, IOENTAILMENT produced the expected entailment result. The case studies therefore provide an executable validation of the uniform encodings developed in this chapter and illustrate how the prover can both expose differences between I/O systems and automatically generate concrete counterexamples for failed entailments.

5.11 Complexity: I/O logics and language fragments

In this section, we use our calculi to give a complete analysis of the computational complexity for deciding entailment across all considered I/O logics. We show that entailment for the full language is co-NP-complete, and investigate two important tractable fragments: Horn conditionals and literal conditionals. Our results are obtained in a uniform way for all considered logics.

Horn (resp. literal) conditionals are pairs in which the input is a conjunction of propositional variables (resp. of *literals*, i.e. propositional variables or their negations) and the output is a propositional variable (resp. a literal). Horn conditionals constitute the foundational fragment of logic programming and rule-based inference [DEGV01]. Literal conditionals play a central role in the causal applications of I/O logics [Boc21]. We show that the complexity of deciding entailment for Horn conditionals is polynomial, and that entailment for literal conditionals exhibits a complexity dichotomy determined by (Or): in the absence of (Or), entailment is polynomial-time decidable, whereas presence of (Or) leads to co-NP-completeness. We also provide a polynomial-time algorithm for the core I/O logics without (Or).

The polynomial-size reductions to classical entailment for the core I/O logics described in Sec. 5.8 (see Tab. 5.3) provide both a co-NP upper bound on complexity (since classical entailment is co-NP-complete) and an efficient method for automation. This bound is tight, since entailment in every considered I/O logic subsumes classical validity, as already observed in [SR17]. Below we provide a simple proof of this result which uses the separation principles.

Lemma 35 ([SR17]). *For every I/O logic $\mathcal{L}$, $(\top, Y)$ is entailed from $\emptyset$ in $\mathcal{L}$ iff Y is classically valid.*

Proof. Follows directly from the separation principles. For $n = 0$ there is only one partition, input sequents (if any) and the ex-falso sequent turn into $\top \vdash \perp$, while the output sequent turns into $\vdash Y$. $\square$

Corollary 3. *Entailment in all considered I/O logics is co-NP-complete.*

For the core I/O logics without (Or), our sequent calculi also support a polynomial-time proof-search procedure for restricted language fragments, where all classical entailments occurring in derivations can be verified in polynomial time. Specifically, to decide whether $G \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\mathcal{L}}$, it suffices to check whether the classical premise of any transition rule application holds (in which case $G \vdash (B, Y)$ is derivable), and whether the classical premise of any pair-elimination rule holds (in which case the problem reduces to a smaller G). There is no need for backtracking, since moving to a premise of a pair-elimination rule never violates derivability, as demonstrated by the following lemma.

Lemma 36 (Semi-invertibility of the pair elimination rules). *Suppose the sequent $G, (A, X) \vdash (B, Y)$ is derivable in $\mathcal{SC}_{\mathcal{L}}$. Then*

- $G \vdash (B, Y \wedge \neg X)$ is derivable in $\mathcal{SC}_{\mathcal{L}}$ if $\mathcal{L}$ does not contain (CT);
- $G \vdash (B \wedge X, Y \wedge \neg X)$ is derivable in $\mathcal{SC}_{\mathcal{L}}$ if $\mathcal{L}$ does contain (CT).

Proof. This fact is a direct consequence of the separation principles: for each partition $(I, J) \in \mathcal{P}(\{1, \dots, |G|\})$ derivability of the required sequents follows from the separation principle for $G, (A, X) \vdash (B, Y)$ for the partition $(I \cup \{|G| + 1\}, J)$ (denoting $(A_{|G|+1}, X_{|G|+1}) = (A, X)$), via trivial equivalences for classical entailment. $\square$

For Horn and literal conditionals, derivability of **LK** premises of the rules in the calculi can be verified in polynomial time using the following simple lemma.

Lemma 37. *Let $\mathcal{A}, \mathcal{B}$ be sets of literals. $\bigwedge_{b \in \mathcal{B}} b$ classically entails $\bigwedge_{a \in \mathcal{A}} a$ iff either $\mathcal{A} \subseteq \mathcal{B}$ or $\{p, \neg p\} \subseteq \mathcal{B}$ for some $p \in \text{Atm}$. $\bigwedge_{b \in \mathcal{B}} b$ classically entails $\bigvee_{a \in \mathcal{A}} a$ iff either $\mathcal{A} \cap \mathcal{B} \neq \emptyset$ or $\{p, \neg p\} \subseteq \mathcal{B}$ or $\{p, \neg p\} \subseteq \mathcal{A}$ for some $p \in \text{Atm}$.*

Proof. Right-to-left implications are trivial; left-to-right implications are proved by contraposition. For the first item, we suppose towards contradiction that $\mathcal{B}$ does not contain conflicting literals (both propositional variable and its negation) and there is a literal $a_0 \in \mathcal{A} \setminus \mathcal{B}$. In such case, $\mathcal{B} \cup \{\bar{a}_0\}$ ⁶ also does not contain conflicting literals, so there exists a valuation on Atm that makes all literals in this set true simultaneously.

⁶Here $\bar{a}$ stands for the dual literal for a , i.e. $\bar{p} = \neg p$ and $\overline{\neg p} = p$ for any $p \in \text{Atm}$.

This valuation makes conjunction $\bigwedge_{b \in \mathcal{B}} b$ evaluate to true and $\bigwedge_{a \in \mathcal{A}} a$ evaluate to false, therefore the assumed classical entailment can not hold. Similarly, for the second item, if neither $\mathcal{A}$ nor $\mathcal{B}$ contain conflicting literals and $\mathcal{A} \cap \mathcal{B} = \emptyset$, then there are no conflicting literals in $\mathcal{B} \cup \{\bar{a} \mid a \in \mathcal{A}\}$, so these literals can be simultaneously turned true by some valuation, which contradicts the classical entailment of $\bigvee_{a \in \mathcal{A}} a$ from $\bigwedge_{b \in \mathcal{B}} b$. $\square$

Consequently, proof search can be performed in polynomial time using the iterative Algorithm 5.1.

Theorem 38. *If the I/O logic $\mathcal{L}$ does not contain (Or) and $G \cup \{(B, Y)\}$ contains only literal conditionals, entailment of (B, Y) from G in $\mathcal{L}$ can be decided in polynomial time via Algorithm 5.1.*

Proof. At each iteration of the **for** loop on line 12, if the condition on line 13 holds, then the current values of G , $\mathcal{B}$, and $\mathcal{Y}$ are replaced with some new values G' , $\mathcal{B}'$, and $\mathcal{Y}'$. Derivability of $G \vdash (\bigwedge_{l \in \mathcal{B}} l, \bigvee_{l \in \mathcal{Y}} l)$ in $\mathcal{SC}_{\mathcal{L}}$ is equivalent to derivability of $G' \vdash (\bigwedge_{l \in \mathcal{B}'} l, \bigvee_{l \in \mathcal{Y}'} l)$ in $\mathcal{SC}_{\mathcal{L}}$: the latter I/O sequent coincides with the second premise of rule (E_1) or (E_3) (depending on whether $\mathcal{L}$ contains (CT)) applied to the former I/O sequent, and the derivability of the **LK** premise of this rule application follows from the condition on the line 13 by Lem. 37, so the derivability of these two I/O sequents is equivalent by this rule application and Lem. 36. Since every such update decreases the size of G , the **repeat** cycle on line 10 is terminated after at most $|G|$ iterations, when no possible application of the pair elimination rule to $G \vdash (\bigwedge_{l \in \mathcal{B}} l, \bigvee_{l \in \mathcal{Y}} l)$ has a derivable **LK** premise. Therefore, after the **repeat** cycle is terminated, derivability of $G \vdash (\bigwedge_{l \in \mathcal{B}} l, \bigvee_{l \in \mathcal{Y}} l)$ in $\mathcal{SC}_{\mathcal{L}}$ is equivalent to derivability of the **LK** premise of at least one of the transition rules of $\mathcal{SC}_{\mathcal{L}}$, which is captured by the expressions on lines 24, 27, and 30 due to Lem. 37. $\square$

Adding the principle (Or) does not affect entailment in the Horn fragment: the set of entailed conditionals is the same both with and without (Or), as demonstrated in the following theorem.

Theorem 39. *Let $\mathcal{L}$ be an I/O logic not containing (Or), and $\mathcal{L}'$ be its extension with (Or). If $G \cup \{(B, Y)\}$ contains only Horn conditionals, then G entails (B, Y) in $\mathcal{L}$ iff G entails (B, Y) in $\mathcal{L}'$.*

Proof. We rely on the following fact about classical entailment: if $B, A_1, \dots, A_m$ are conjunctions of propositional variables and B classically entails $\bigvee_{i=1}^m A_i$ (equivalently, $B \vdash A_1, \dots, A_m$ is derivable in **LK**) then B classically entails A_k (equivalently, $B \vdash A_k$ is derivable in **LK**) for some $k \in \{1, \dots, m\}$. It is true, since otherwise $B \vdash A_1, \dots, A_m$ is falsified by boolean assignment ρ defined by $\rho(x) = \top$ iff $x \in B$.

Only the left-to-right direction of the theorem's claim is non-trivial, and it follows by a simple analysis of the separation principles: for $\mathcal{L}' = \mathbf{OUT}_2^+$ and $\mathcal{L}' = \mathbf{OUT}_4^+$ derivability

Algorithm 5.1: Algorithm for deciding if G entails $(\mathcal{B}, y)$ in an I/O logic $\mathcal{L}$ without (OR) for the literal fragment. Literal conditionals are represented as pairs $(\mathcal{A}, x)$ where $\mathcal{A}$ is a set of literals, and x is a literal. neg is the operation negating a literal (i.e. $neg(p) = \neg p$ and $neg(\neg p) = p$ for every $p \in Atm$).

```
1 Function Contradictory( $\mathcal{A}$ ):
2   for  $x \in \mathcal{A}$  do
3     if  $neg(x) \in \mathcal{A}$  then
4       return true
5     end
6   end
7   return false
8 Function Check_Entailment_Literal_noOr( $\mathcal{L}, G, \mathcal{B}, y$ ):
9    $\mathcal{Y} \leftarrow \{y\}$ 
10  repeat
11     $updated \leftarrow false$ 
12    for  $(\mathcal{A}, x) \in G$  do
13      if  $\mathcal{A} \subseteq \mathcal{B}$  or Contradictory( $\mathcal{B}$ ) then
14         $\mathcal{Y} \leftarrow \mathcal{Y} \cup \{neg(x)\}$ 
15        if  $\mathcal{L} \in \{OUT_3, OUT_3^\perp, OUT_3^+\}$  then
16           $\mathcal{B} \leftarrow \mathcal{B} \cup \{x\}$ 
17        end
18         $G \leftarrow G \setminus \{(\mathcal{A}, x)\}$ 
19         $updated \leftarrow true$ 
20      end
21    end
22  until not  $updated$ 
23  if  $\mathcal{L} \in \{OUT_1, OUT_3\}$  then
24    return Contradictory( $\mathcal{Y}$ )
25  end
26  if  $\mathcal{L} \in \{OUT_1^\perp, OUT_3^\perp\}$  then
27    return Contradictory( $\mathcal{B}$ ) or Contradictory( $\mathcal{Y}$ )
28  end
29  if  $\mathcal{L} \in \{OUT_1^+, OUT_3^+\}$  then
30    return  $\mathcal{B} \cap \mathcal{Y} \neq \emptyset$  or Contradictory( $\mathcal{B}$ ) or Contradictory( $\mathcal{Y}$ )
31  end
```

of the output sequent of $\mathcal{L}'$ implies derivability of either the output sequent or one of the input sequents of $\mathcal{L}$, and in all other cases the output sequents of $\mathcal{L}$ and $\mathcal{L}'$ are the same; for all logics, derivability of the input sequent of $\mathcal{L}'$ for $I \neq \emptyset$ implies derivability of one of the input sequents of $\mathcal{L}$, and derivability of the input sequent of $\mathcal{L}'$ for $I = \emptyset$ implies derivability of the ex-falso sequent of $\mathcal{L}$. $\square$

Therefore, entailment for Horn conditionals can be decided in polynomial time using the same Algorithm 5.1.

In contrast, for the literal fragment, adding the principle (Or) not only enlarges the set of entailed conditionals (e.g., for $a, b \in \text{Atm}$, $\{(a, b), (\neg a, b)\}$ entails $(\top, b)$ in I/O logics with (Or), but not in I/O logics without (Or)), but also raises the computational complexity to co-NP, matching that of the full language.

Lemma 38. *If the I/O logic $\mathcal{L}$ contains the principle (Or) and $G \cup \{(B, Y)\}$ only contains literal conditionals, deciding whether G entails (B, Y) in $\mathcal{L}$ is co-NP-hard.*

Proof. We prove co-NP-hardness by reducing classical validity of a formula in Disjunctive Normal Form (DNF) to entailment in $\mathcal{L}$. Let $F = \bigvee_{i=1}^n \bigwedge_{j=1}^{m_i} l_{ij}$ be a formula in DNF (i.e. each l_{ij} is a literal). For an arbitrary propositional variable x not appearing in F , consider the following set of literal conditionals corresponding to F : $G(F) = \{(\bigwedge_{j=1}^{m_i} l_{ij}, x)\}_{i=1}^n$. Classical validity of F is equivalent to entailment of $(\top, x)$ from $G(F)$ in $\mathcal{L}$, which can be seen by simple inspection of the separation principles for these logics. Namely, for $J \neq \emptyset$ the output sequent is trivially derivable, and for $J = \emptyset$ either the output sequent is $\top \vdash x, \bigwedge_{j=1}^{m_1} l_{1j}, \dots, \bigwedge_{j=1}^{m_n} l_{nj}$ and there is no input sequents (in cases $\mathcal{L} = \mathbf{OUT}_2^+$ and $\mathcal{L} = \mathbf{OUT}_4^+$), or the output sequent is $\vdash x$ and the input sequent is $\top \vdash \bigwedge_{j=1}^{m_1} l_{1j}, \dots, \bigwedge_{j=1}^{m_n} l_{nj}$. So in both cases derivability of one of the sequents for $J = \emptyset$ (and thus entailment of $(\top, x)$ from $G(F)$) is equivalent to classical validity of F . $\square$

Consequently, for I/O logics, the complexity of entailment over the literal fragment is either PTIME or co-NP, depending solely on the presence of (Or) in the logic.

5.12 Complexity: Bochman's actual cause is in NP

Causal I/O logics provide a formal basis for reasoning about causality, and the methods presented in the previous section can be used to support their automated reasoning. In particular, [Boc04] introduced two non-monotonic semantics for causal theories represented as conditionals –called *general* and *causal* non-monotonic semantics– capturing the “*principles of explanation closure*” and the “*principle of universal causation*”, respectively. A central notion in this framework is strong equivalence: strongly equivalent theories are interchangeable in any larger causal theory without affecting the resulting non-monotonic semantics. It was shown that two causal theories are strongly equivalent with respect to the general (resp. causal) non-monotonic semantics iff they entail the

same set of conditionals in $\mathbf{OUT}_3^\perp$ (resp. $\mathbf{OUT}_4^\perp$). This can be verified by checking mutual entailment of their conditionals in the corresponding logic, using the propositional encodings in Tab. 5.3, or, for literal causal theories (introduced in [Boc21] and used in most applications), by the Algorithm 5.1.

Beyond theory comparison, entailment in $\mathbf{OUT}_3^\perp$ was used in [Boc18] to define a notion of *actual cause* for literal causal theories⁷, providing an alternative to the standard definition of Halpern and Pearl [HP05]. We state the formal definition below and refer to [Boc18] for the motivation and conceptual analysis of this notion.

Definition 31 ([Boc18]). *For a given G containing only literal conditionals, Boolean valuation ν , literals l_0 and l_1 true w.r.t. ν , l_0 is called actual cause of l_1 in ν w.r.t. G iff there exists set L of literals true in ν , such that G entails $(l_0 \wedge \bigwedge_{l \in L} l, l_1)$, but does not entail $(\bigwedge_{l \in L} l, l_1)$ in $\mathbf{OUT}_3^\perp$.*

Since, as we have shown, entailment from a literal causal theory in $\mathbf{OUT}_3^\perp$ is decidable in polynomial time, actual causation can be checked in nondeterministic polynomial time by guessing the set L (considering only variables appearing in G , l_0 , and l_1 is sufficient) and verifying the two entailment conditions. As a result, actual causation under Bochman’s definition is decidable in NP, making it computationally simpler than the Halpern–Pearl definition, whose verification is Δ_2^P -complete [EL02]. Establishing NP-hardness for Bochman’s notion remains open, suggesting that its exact complexity may be even lower.

⁷It is shown there how the translation from [BL15] can be used to represent many standard examples from the causality literature—originally expressed using structural equation models—via literal causal theories.

Part III

Typicality

Locally non-monotonic reasoning

Conditional logics are tightly connected with a major line of research in symbolic AI: non-monotonic reasoning. Non-monotonic reasoning aims to capture defeasible forms of inference, where conclusions may need to be withdrawn in the presence of additional evidence or exceptional circumstances. This behavior is naturally captured through the typicality interpretation of conditionals, where a statement of the form $(A \rightsquigarrow B)$ is understood to mean that “if A holds, then normally/typically B follows”, while allowing for exceptional cases where A holds but B does not. A standard example is the conditional $(\text{bird} \rightsquigarrow \text{flies})$, expressing that birds typically fly, though some exceptional birds—such as penguins—may fail to do so.

Formally, when dealing with the entailment of conditionals, one can distinguish between two different dimensions of non-monotonicity, as observed in [Boc06]. The first is *local* non-monotonicity: an entailment of a conditional $(A \rightsquigarrow B)$ from a set of conditionals G may cease to hold when the antecedent A is strengthened, reflecting the possibility of exceptions to more general rules. The second is *global* non-monotonicity: an entailment of a conditional $(A \rightsquigarrow B)$ from a set of conditionals G may cease to hold when G is extended with additional conditionals, reflecting revision of conclusions in light of new information. Preferential semantics has been established as the core model-theoretic framework for defining, studying, and classifying formalisms that exhibit both forms of non-monotonicity [Sho87, KLM90, LM92]. In the following two chapters, we investigate both dimensions of non-monotonic conditional entailment through their central preference-based formalisms: this chapter addresses the KLM hierarchy of logics, a fundamental framework that isolates the formal properties of locally non-monotonic (yet globally monotonic) reasoning and their realization via preferential semantics; the subsequent chapter turns to Rational Closure, the canonical extension of the KLM paradigm designed to obtain a conditional entailment operator more suitable for practical applications, and exhibiting global non-monotonicity in particular.

Introduced by Kraus, Lehmann, and Magidor [KLM90], the framework of KLM logics defines a hierarchy of inference relations — **C**, **CL**, **P**, **CM**, **M** — characterized both axiomatically and semantically, bringing together abstract principles of locally non-monotonic inference with preferential semantics based on the notion of “more normal” situations. The resulting hierarchy ranges from weak cumulative inference to fully monotonic entailment, with the core logic **P** occupying a central position: its axioms, now widely referred to as the KLM principles, have become a standard benchmark for evaluating defeasible inference systems.

This chapter starts with a historical background (Sec. 6.1) and state of the art regarding analytic characterizations and automated deduction in KLM logics (Sec. 6.2). After that, KLM hierarchy is introduced formally, both axiomatically and via preferential semantics (Sec. 6.3). We then begin our development by introducing an “analytic” version of the smoothness property for preferential models, yielding an alternative characterization of the weakest KLM logic **C**, which is easier to satisfy and which exhibits better computational behaviour (Sec. 6.4). Building on this, we develop new analytic sequent calculi for the entire KLM hierarchy (Sec. 6.5), where derivations proceed via saturation of the graph of conditional dependencies between formulas occurring in the entailment. Central to this approach is the notion of a separating graph, which can be seen as a generalization of the separation principles introduced for I/O logics in Chapter 5. Using these characterizations, we obtain a complete classification of the computational complexity of all KLM logics — resolving coNP-completeness of **C** — as well as of their Horn and literal fragments (Sec. 6.6). Finally, we present the automated prover `KLMP` `propagator` [CER24], relying on our characterizations (Sec. 6.7).

6.1 Background

The search for formal systems capable of representing defeasible commonsense reasoning became one of the major themes of symbolic AI in the 1980s. A variety of different formalisms for non-monotonic reasoning were proposed during this period. Reiter’s default logic [Rei80] approached non-monotonicity procedurally via rule-based derivations with consistency checks. McCarthy’s circumscription [McC80] modeled defeasible reasoning semantically via minimization of abnormality predicates. Adams developed a probabilistic account of defeasible conditionals based on high conditional probabilities [Ada75]. Moore’s autoepistemic logic [Moo85] represented non-monotonic reasoning as a form of modal (epistemic) reasoning about an agent’s own beliefs. These approaches were originally developed largely independently and differed significantly both technically and conceptually, motivating the search for a more abstract and fundamental understanding of non-monotonic inference itself. On the syntactic side, Gabbay [Gab85] proposed general “minimal” postulates for non-monotonic consequence relation, later developed into a systematic abstract theory by Makinson [Mak89]. In parallel, Shoham [Sho87] proposed to use preferential models as a unifying semantic framework for non-monotonic reasoning, based on selecting the “most normal” or “preferred” situations.

These strands were brought together in the milestone work of Kraus, Lehmann, and Magidor [KLM90], which introduced a hierarchy of both monotonic and non-monotonic logics now known as the KLM hierarchy. The paper systematically characterized several systems both axiomatically and semantically via preferential models. The hierarchy includes five logics: logic **C** of cumulative reasoning, which corresponds to the system studied by Gabbay and Makinson; logic **CL** of cumulative reasoning with Loop, which extends **C** to work with more natural class of preferential models; logic **P** of preferential reasoning, which corresponds directly both to Shoham’s semantical proposition and Adam’s probabilistic system; logic **CM** of cumulative monotonic reasoning, which is monotonic yet weaker than classical logic; and finally logic **M**, which coincides with classical propositional logic. Although KLM hierarchy was described in terms of consequence relations rather than conditionals, a close connection to conditional logic was clear from the very beginning: KLM logics can be seen as the flat fragment of conditional logics (i.e. without nesting of conditionals), where **P** in particular coincides with the flat fragment of both **PCL** from Chapter 3 and Åqvist’s **F_{CM}** from Chapter 4.

The KLM framework had a foundational impact on the study of defeasible reasoning, and the axioms of system **P** became widely known as the “KLM postulates,” serving as a standard benchmark against which non-monotonic systems are evaluated. The framework was subsequently developed and generalized in several directions, most notably by Arieli and Avron [AA00], who extended it to multiple-conclusion and non-classical consequence relations, thereby accommodating paraconsistent reasoning. It also served as the basis for major later developments in defeasible reasoning, including ranked approaches such as Pearl’s System *Z* [Pea90] and rational closure [LM92], and established important connections with other areas of symbolic AI, such as belief revision [KS91].

6.2 State of the art

The central logic **P** of KLM framework has been extensively studied from different perspectives, including the analytic one. A number of analytic calculi have been developed for **P**, both directly based on the semantics of **P** [GGOP07, GGOP09, BV18] or covering it as a fragment of more expressive conditional formalisms [Bou90, BHM09]. On the semantic side, small-model constructions were suggested for **P** to establish its coNP-completeness [LM92]; these constructions were later generalized to nested conditional case by Friedman and Halpern [FH94] into the modular framework discussed in Chapters 3 and 4.

The remaining non-classical KLM logics have received considerably less attention. A major step toward their uniform analytic treatment was made in [GGOP09], where tableau calculi for all non-monotonic logics KLM (and some related variants) was introduced by enriching the language with a **GL**-style modality encoding the underlying preference relation. This original approach underlies our nested-case developments in Chapters 3 and 4. It was used as a base for the automated prover *KLMLean* [GGP07] and as a formal tool to establish coNP-completeness for all covered logics except for **C**, for which

the calculi required explicit loop checks and yielded only a coNEXP upper bound. It was improved in [AOP16], which introduced nested sequent calculi for a number of normal conditional logics, including the basic logic **CK**, whose flat coincides with **C**. Their approach lowered the upper complexity bound for **C** to PSPACE and led to a more efficient automated proof search, implemented in the prover NESCOND [OP14], though it is still showed worse performance than for other KLM logics and the exact complexity remained out of reach. This difficulty to obtain well-behaved analytic characterization for **C** can be traced to its semantics: unlike the stronger KLM systems, its preferential models do not assume the preference relation to be a partial order, while still excluding infinite descending chains. As already observed in the original KLM paper [KLM90, p.21], such models are difficult to construct and manipulate, making this semantics a poor guide for analytic proof theory.

6.3 Preliminaries: KLM hierarchy

[KLM90], henceforth referred to as “the KLM paper”, introduced axiomatizations for the binary relation $\sim$ (interpreted as “... typically implies ...”) along with appropriate preferential semantics. Henceforth we will use the notation $(A \rightsquigarrow B)$ (instead of $A \sim B$), for the non-monotonic *conditionals*, where A (*condition*) and B (*consequence*) are propositional formulas of classical logic (called *components* of the conditional). We will denote classical entailment from a set of propositional formulas Γ with $Cn(\Gamma)$ and classical equivalence between two formulas with $\equiv_{Cn}$. On the syntactic side, the logics of the KLM family are defined via the basic inference rules, which we will refer to as *original KLM rules*, to distinguish them from the modified systems of rules that we will suggest later. Different KLM logics are given by different subsets of these rules. The basic logic **C** of cumulative reasoning consists of the following five base rules:

- (Id) $(A \rightsquigarrow A)$ is derivable from no premises
- (RWk) $(A \rightsquigarrow B)$ derives $(A \rightsquigarrow B')$ whenever $B' \in Cn(\{B\})$
- (LEq) $(A \rightsquigarrow B)$ derives $(A' \rightsquigarrow B)$ whenever $A \equiv_{Cn} A'$
- (CMon) $(A \rightsquigarrow B)$ and $(A \rightsquigarrow C)$ derive $(A \wedge B \rightsquigarrow C)$
- (CCut) $(A \rightsquigarrow B)$ and $(A \wedge B \rightsquigarrow C)$ derive $(A \rightsquigarrow C)$

The central of them is the “cautious monotonicity” rule (CMon) introduced by Gabbay as restricted form of strengthening of antecedent for non-monotonic inference systems [Gab85]. Other KLM logics extend this base with some of the following additional rules:

- (Loop) $(A_1 \rightsquigarrow A_2), \dots, (A_{n-1} \rightsquigarrow A_n), (A_n \rightsquigarrow A_1)$ derive $(A_1 \rightsquigarrow A_n)$
- (Or) $(A_1 \rightsquigarrow B)$ and $(A_2 \rightsquigarrow B)$ derive $((A_1 \vee A_2) \rightsquigarrow B)$

Logic	Defining inference principles
C	(Id), (RWk), (LEq), (CMon), (CCut)
CL	(Id), (RWk), (LEq), (CMon), (CCut), (Loop)
P	(Id), (RWk), (LEq), (CMon), (CCut), (Or)
CM	(Id), (RWk), (LEq), (CMon), (CCut), (Mon)
M	(Id), (RWk), (LEq), (CMon), (CCut), (Or), (Mon)

Table 6.1: The axiomatic definition of KLM logics

(Mon) $(A \rightsquigarrow B)$ derives $(A' \rightsquigarrow B)$ whenever $A \in Cn(\{A'\})$

The logic **CL** extends **C** with the rule (Loop), and the logic **P** of preferential reasoning, extends **C** with the rule (Or) for reasoning by cases. The KLM family also includes two monotonic logics, the logic **CM** of cumulative monotonic reasoning, which extends **C** with (Mon), and the logic of monotonic reasoning **M** which extends **CM** with (Or). Logic **M** collapses to classical logic [KLM90], while **CM** is strictly weaker. Notice that both **P** and **CM** are stronger than **CL**, as they derive the rule (Loop).

Notice that there is a significant intersection between these inference principles and those used in the previous chapter to define I/O logics (Sec. 5.3), although different names are used. Moreover, **CM** actually coincides with the throughput logic $\mathbf{OUT}_3^+$: their axiomatization differ only by rule (*LEq*) in **CM** and (And) in $\mathbf{OUT}_3^+$, both of which can be derived from the remaining rules [KLM90]. Similarly, as already noted in Chapter 5, both $\mathbf{OUT}_2^+$ and $\mathbf{OUT}_4^+$ were shown to collapse to classical logic [MvdT01], and hence coincide with the KLM logic **M**. By contrast, the non-monotonic members of the KLM hierarchy crucially lack unrestricted strengthening of the antecedent, expressed by (SI) in I/O logic and (*Mon*) in the KLM framework.

Semantically, KLM logics are characterized by another variant of preference models. Similarly to deontic preference models in Chapter 4 (Def. 7,p.49), these models satisfy absoluteness, i.e. have one preference relation that is not world-dependent. This kind of models, add one generalization w.r.t. preferential models of Chapter 3 and Chapter 4: each world (called ‘state’ in the KLM context) can be labeled not with exactly one, but with an arbitrary number of Boolean valuations. This generalization allows to cover the logics that do not satisfy (Or).

Definition 32. A (KLM) preferential model is a triple $\langle \mathcal{S}, \preceq, l \rangle$ where $\mathcal{S}$ is a set of abstract states, l is a labeling function that maps each state in $\mathcal{S}$ to a (possibly empty) set of Boolean assignments on the set of propositional atoms Atm , the preference relation $\preceq$ is a binary relation on $\mathcal{S}$. The strict version of $\preceq$, is defined as follows: $v \prec u$ when $v \preceq u$ and $u \not\preceq v$.

The satisfaction of a formula in a state s is interpreted as satisfaction by all boolean assignments in $l(s)$, and the satisfaction of a conditional is defined as usual, although traditionally using minimal states of maximal ones.

Logic	Conditions on preference relation $\succeq$	Conditions on labeling function l
C	smooth	-
CL	smooth and transitive	-
P	smooth and transitive	functional
CM	empty	-
M	empty	functional

Table 6.2: Conditions on preference models for KLM logics

Definition 33. A formula A is satisfied in a state s of a model $M = \langle \mathcal{S}, \succeq, l \rangle$, (in symbols $M, s \models A$) if it is satisfied by all Boolean assignments $l(s)$. We denote by $\|A\|^M$ the truth set of A , i.e., the subset of states in $\mathcal{S}$ that satisfy A . A conditional $(A \rightsquigarrow B)$ is satisfied by the model M (in symbols $M \models (A \rightsquigarrow B)$) if $M, s \models B$ for all $s \in \min_{\prec}(\|A\|^M)$ where $\min_{\prec}(X) = \{s \in X : \nexists s' \in X \text{ s.t. } s' \prec s\}$.

A truth set of a formula may have multiple minimal states, but the familiar condition of *smoothness*, which already appeared in the context of Åqvist's logics (Sec. 4.3) is also used in KLM logics to ensure that all truth sets have at least one minimal state.

Definition 34. A KLM preferential model is smooth if for any formula A and state $s \in \|A\|^M$, there is $s' \in \min_{\prec}(\|A\|^M)$ such that either $s' = s$ or $s' \prec s$.

Smoothness is the defining feature of the basic logic **C**. However, as already noted in [KLM90], working with it can be challenging due to the difficulty of checking this condition for all truth sets. Additional conditions on preference models characterize the other KLM logics (see Tab. 6.2). In particular, **CL** and **P** assume the preference relation to be transitive, and for the monotonic KLM logics **CM** and **M** the preference relation is empty. Additionally, for the logics **P** and **M**, the labeling has exactly one Boolean assignment for each state (we call this property *functionality* of a labeling function).

In the following, when discussing a KLM logic $\mathcal{L} \in \{\mathbf{C}, \mathbf{CL}, \mathbf{CM}, \mathbf{P}, \mathbf{M}\}$, we will refer to the models satisfying the respective properties as $\mathcal{L}$ -models.

Theorem 40. [KLM90] A set Γ of conditionals entails a conditional $\mathcal{E}$ in a KLM logic $\mathcal{L}$ (in symbols, $\Gamma \vdash_{\mathcal{L}} \mathcal{E}$) iff every $\mathcal{L}$ -model that satisfies all conditionals in Γ satisfies also $\mathcal{E}$.

6.4 Analytic version of smoothness condition

In the KLM paper, the authors show [KLM90, Lem.11 and Lem.17] that their canonical models for **C** and **CL** also satisfy a stronger version of smoothness that requires each non-empty truth set to have the least state, i.e. the state that is less preferable than all other states in the set. We will refer to this condition as *l-smoothness*.

Definition 35. *The set of the least states in a model $M = \langle \mathcal{S}, \preceq, l \rangle$ is defined as follows: $\text{least}_{\prec}(X) = \{s \in X \mid s \prec x \text{ for all } x \in X \setminus \{s\}\}$. M is called *l-smooth* if for any formula A , $\|A\|^M \neq \emptyset$ implies $\text{least}_{\prec}(\|A\|^M) \neq \emptyset$.*

Although the canonical models for the other KLM logics do not satisfy l-smoothness, it is easy to see that entailment in **CM** and **M** can be characterized by models with exactly one state (and, hence, l-smooth): since the preference relation is empty, in a countermodel for entailment $\Gamma \vdash_{\mathcal{L}} \mathcal{E}$ only the state that falsifies the conditional $\mathcal{E}$ is needed. Moreover, as shown in [LM92, Lem.8] and also small model constructions **PCLTA** and Åqvist's **F_{CM}**, *finite* entailments in **P** can be characterized by finite linearly-ordered (and, hence, l-smooth) preference models. In sum, in the semantical characterizations of finite entailment, smoothness can be replaced by l-smoothness for all KLM logics. As we only deal with finite entailments, we will use both versions of smoothness.

Neither smoothness nor l-smoothness is a frame property, since both depend on the valuation (labeling) in the model. This makes them difficult to capture analytically, as one has to ensure the corresponding conditions for arbitrary formulas, rather than only for the formulas appearing in the entailment under consideration. For KLM logics stronger than **C**, this issue is bypassed by the fact that, over finite models, smoothness is implied by the transitivity conditions imposed on the preference relation. For **C**, however, no similarly simple frame condition is known, even in the case of finite models. To overcome this difficulty, we introduce the following *analytic* version of l-smoothness relative to a given set of input conditionals, restricting the condition to input antecedents. We call this condition *analytic* because it is determined entirely by the formulas occurring in the entailment under consideration: unlike l-smoothness, which imposes a requirement on the truth set of every formula in the language, it only needs to be checked for the finitely many antecedents occurring in the input.

Definition 36. *$M = \langle \mathcal{S}, \preceq, l \rangle$ satisfies analytic l-smoothness w.r.t. finite set of conditionals $\{(A_i \rightsquigarrow B_i)\}_{i=0}^n$ when $\|A_i\|^M \neq \emptyset$ implies $\text{least}_{\prec}(\|A_i\|^M) \neq \emptyset$ for every $i \in \{0, \dots, n\}$.*

This condition is weaker than l-smoothness, but we can show that it provides an adequate characterization for *finite* entailments in **C**.

Theorem 41. *For a finite Γ , $\Gamma \vdash_{\mathbf{C}} \mathcal{E}$ iff there is no model M satisfying analytic smoothness w.r.t. $\Gamma \cup \{\mathcal{E}\}$, such that $M \not\models \mathcal{E}$ and $M \models \mathcal{C}$ for each $\mathcal{C} \in \Gamma$.*

Proof. Since analytic l-smoothness is implied by l-smoothness, which characterizes finite entailment in **C**, it suffices to show that any model M from the theorem statement can be turned into an l-smooth countermodel. Without loss of generality, we restrict the set of atomic variables Atm to the variables occurring in $\Gamma \vdash \mathcal{E}$ (any l-smooth countermodel can be extended to any larger set of variables by assigning the additional variables to false in all labelings).

Consider the following model transformation $M \mapsto M^*$ “repairing” l-smoothness for one formula A : if M does not satisfy l-smoothness for A (so both $\|A\|^M \neq \emptyset$ and $\text{least}_{\prec}(\|A\|^M) = \emptyset$), $M^* = \langle \mathcal{S}^*, \preceq^*, l^* \rangle$ extends $M = \langle \mathcal{S}, \preceq, l \rangle$ with one additional state s^* , which we make the least state in $\|A\|^{M^*}$:

- $\mathcal{S}^* = \mathcal{S} \cup \{s^*\}$,
- $s^* \preceq^* s$ when $s \in \|A\|^M$,
- $s \preceq^* s^*$ when $s \notin \|A\|^M$,
- $s \preceq^* s'$ iff $s \preceq s'$ for any $s, s' \in \mathcal{S}$,
- $l^*(s^*) = \{\rho \mid \rho \models A\}$,
- $l^*(s) = l(s)$ for any $s \in \mathcal{S}$.

We have $s^* \in \text{least}_{\prec}(\|A\|^{M^*})$ by definition. We show that the transformation also preserves l-smoothness for any $B \neq A$. If $\|B\|^M = \emptyset$ then $B \notin \text{Cn}(A)$ (since $\|A\|^M \neq \emptyset$), so $l^*(s^*) \not\models B$, so $\|B\|^{M^*} = \emptyset$. If $\|B\|^M \neq \emptyset$ and $s_B \in \text{least}_{\prec}(\|B\|^M)$, s_B remains the least after the model transformation: it can only be violated if both $l^*(s^*) \models B$ and $s_B \not\preceq^* s^*$ in M^* , but the former implies $B \in \text{Cn}(A)$ and the latter implies $s_B \in \|A\|^M$, which would imply $s_B \in \text{least}_{\prec}(\|A\|^M)$ contradicting the absence of l-smoothness on A in M .

Hence any countermodel M satisfying analytic l-smoothness for $\Gamma \cup \{\mathcal{E}\}$ can be converted into an l-smooth countermodel in a finite number of steps: select a maximal set of pairwise classically non-equivalent formulas (which is finite, given that so is Atm) and “repair” l-smoothness for each one using the transformation above. The resulting model achieves l-smoothness for all formulas in the language while still satisfying Γ and falsifying $\mathcal{E}$ (since the transformation preserves the least states). $\square$

Remark 12. *Analytic l-smoothness of M w.r.t. set $\mathcal{A}$ can be verified in polynomial time w.r.t. $|M|$ and cumulative size of $\mathcal{A}$, and as we will show in the completeness proof for the calculus for $\mathbf{C}$, a countermodel with analytic smoothness of a polynomial size exists for every finite entailment. In contrast, it is unclear whether smoothness (or l-smoothness) from the KLM paper always permits a polynomial-size countermodel, and whether these properties (which apply to all formulas in the language) can be verified in polynomial time.*

6.5 Saturating sequent calculi for KLM logics

In this section, we present analytic sequent calculi for the KLM logics. As in the calculi for I/O logics, our calculi manipulate two kinds of sequents: conditional sequents of the form $\Gamma \vdash \mathcal{E}$, where Γ is a finite set of conditionals, and $\mathcal{E}$ a conditional, and standard **LK**

Concluding rule: $\frac{}{\Pi, \mathcal{E} \vdash \mathcal{E}}$

Saturating rule form: $\frac{\{\Pi, \Sigma, \delta_j \vdash \mathcal{E}\}_{j=1}^m \quad \mathfrak{G}}{\Pi, \Sigma \vdash \mathcal{E}}$

Rule	m	Σ	$\{\delta_j\}_{j=1}^m$	$\mathfrak{G}$
(Id^s)	$m = 1$	$\emptyset$	$\{A \rightsquigarrow A\}$	no condition
$(Equiv^s)$	$m = 1$	$\{A \rightsquigarrow A', A' \rightsquigarrow A, A \rightsquigarrow B\}$	$\{A' \rightsquigarrow B\}$	no condition
$(Loop^s)$	$m = 1$	$\{A_i \rightsquigarrow A_{i+1}, \dots, A_n \rightsquigarrow A_1\}$	$\{A_1 \rightsquigarrow A_n\}$	no condition
(Tr^s)	$m = 1$	$\{A \rightsquigarrow B, B \rightsquigarrow C\}$	$\{A \rightsquigarrow C\}$	no condition
(Cn^s)	$m = 1$	$\{A \rightsquigarrow B_i\}_{i=1}^n$	$\{A \rightsquigarrow C\}$	$\{B_i\}_{i=1}^n \vdash C$
(CnM^s)	$m \geq 1$	$\{A \rightsquigarrow B_i\}_{i=1}^n$	$\{A \rightsquigarrow C_j\}_{j=1}^m$	$\{B_i\}_{i=1}^n \vdash \{C_j\}_{j=1}^m$

Analyticity restriction: $\mathcal{F}(\delta_j) \subseteq \mathcal{F}(\Pi \cup \Sigma)$

Novelty restriction: $\delta_j \notin \Pi \cup \Sigma$

Figure 6.1: Saturating calculi for KLM logics. $\mathcal{F}(\Gamma)$ denotes the set of all components of conditionals in Γ .

sequents $A_1, \dots, A_n \vdash B_1, \dots, B_m$ where $A_1, \dots, A_n, B_1, \dots, B_m$ are Boolean formulas. The rules are presented in Fig. 6.1 and their intuition is explained in the respective sections dedicated to each logic. The superscript s on the rule's names serves to distinguish them from the corresponding KLM principle (if any). Our calculi are constructed using a modular approach, akin to the calculi in the KLM paper. The calculus $\mathcal{SC}_{\mathbf{C}}$ for the basic logic $\mathbf{C}$ consists of the rules (Id^s) , $(Equiv^s)$, and (Cn^s) . The extension $\mathcal{SC}_{\mathbf{CL}}$ for $\mathbf{CL}$ includes also the rule $(Loop^s)$. The calculi $\mathcal{SC}_{\mathbf{CM}}$, $\mathcal{SC}_{\mathbf{P}}$ and $\mathcal{SC}_{\mathbf{M}}$ for $\mathbf{CM}$, $\mathbf{P}$, and $\mathbf{M}$ are achieved by adding the rules (Tr^s) , (CnM^s) , and both, respectively.

All rules follow a specific structure: the sequent(s) in the premise(s) and in the conclusion differ only for the presence in the former of one conditional δ_j on the left hand side (LHS). The rules (Cn^s) and (CnM^s) additionally contain an **LK** premise $\mathfrak{G}$. Note that all rules, except (CnM^s) , have only one conditional sequent among premises.

In the bottom-up proof search for a sequent $\Gamma \vdash \mathcal{E}$, the set of conditionals on the LHS expands. A branch is closed by the concluding rule when $\mathcal{E}$ appears on the LHS.

Crucially, our rules in Fig. 6.1 apply only under two restrictions: (1) the conditional δ_j is not already on the LHS of the rule premise (*novelty*), and (2) the components in δ_j are already present among the components of the conditionals in the conclusion (*analyticity*). Note that our analyticity condition is, in a sense, stronger than the subformula property, which (only) requires that all formulas in the premises of a rule are subformulas of formulas in the conclusion. The following graph view of derivations in our calculi, clarifies this matter.

Graphs of conditional dependencies offer a convenient representation of $\mathcal{SC}_{\mathcal{L}}$ derivations and proofs. In this perspective we view the set of formulas from all components of conditionals in $\Gamma \cup \mathcal{E}$ (in symbols, $\mathcal{F}(\Gamma \cup \{\mathcal{E}\})$) as nodes in a graph, and conditionals on

the LHS as directed edges between them. Each one-premise rule adds the edge δ_1 to the current graph, while (CnM^s) creates m copies of the graph with one additional edge $\{\delta_j\}_{j=1}^m$ each.

Definition 37 (Graph view). *A derivation in $\mathcal{SC}_{\mathcal{L}}$ of a sequent $\Gamma \vdash \mathcal{E}$ is the process of saturating a graph G (graphs, in case of $\mathcal{SC}_{\mathcal{P}}$ and $\mathcal{SC}_{\mathcal{M}}$) with nodes $V = \mathcal{F}(\Gamma \cup \{\mathcal{E}\})$ and directed edges $E(G) \supseteq \Gamma$. Starting from the graph with Γ as the set of edges, each application of a rule in $\mathcal{SC}_{\mathcal{L}}$ to a graph with edges E replaces this graph with m graphs $\{(V, E \cup \{\delta_j\})\}_{j=1}^m$. A proof in $\mathcal{SC}_{\mathcal{L}}$ is a derivation in which an edge corresponding to $\mathcal{E}$ appears in each graph. at the end of the process. If this condition remains unmet despite exhaustive rule applications, the derivation is said to be failed.*

The analyticity condition prevents the expansion of nodes from $V = \mathcal{F}(\Gamma \cup \{\mathcal{E}\})$, and the novelty condition prohibits the addition of duplicate directed edges between the same pair of nodes. This results in the process reaching saturation – where no further edges can be added to the graphs by applying the calculus rules – within a polynomial (at most $(2 \cdot |\Gamma| + 2)^2$) number of rule applications in each branch.

Notice that *saturation* w.r.t. a single-premise rule (r) means that regardless of the choice of Σ among the edges of the graph, the edge δ_1 is already in the graph. Consequently, (r) cannot be applied due to the novelty restriction. For the rule (CnM^s) saturation is achieved if at least one of the edges $\{\delta_j\}_{j=1}^m$ is already present in the graph. A graph with nodes $\mathcal{F}(\Gamma \cup \{\mathcal{E}\})$ and starting edges Γ , reaching saturation w.r.t. all rules of $\mathcal{SC}_{\mathcal{L}}$ only, corresponds to either a proof of $\Gamma \vdash \mathcal{E}$ or a failed derivation. The latter provide a counterexample to the derivability of $\Gamma \vdash \mathcal{E}$.

However, to prove the completeness of our calculi and allow them to be encoded into SMT solvers, we will generalize the notion of counterexample to *separating graphs*. A separating graph for $\Gamma \vdash \mathcal{E}$ contains all conditionals from Γ as edges, but no edge $\mathcal{E}$, while still being saturated w.r.t. the rules of $\mathcal{SC}_{\mathcal{L}}$.

Definition 38. $G = (V(G), E(G))$ is a separating graph for $\Gamma \vdash \mathcal{E}$ w.r.t. $\mathcal{SC}_{\mathcal{L}}$ if it satisfies the following four conditions: (1) $V(G) = \mathcal{F}(\Gamma \cup \{\mathcal{E}\})$, (2) $\Gamma \subseteq E(G)$, (3) $\mathcal{E} \notin E(G)$, and (4) for every rule (instance) of $\mathcal{SC}_{\mathcal{L}}$, if $\Sigma \subseteq E(G)$ then at least one of the edges $\{\delta_j\}_{j=1}^m$ belongs to $E(G)$.

Although the edges in a separating graph for $\Gamma \vdash \mathcal{E}$ are not required to be derivable from Γ , all edges *not present* in the graph are *not* derivable from Γ due to saturation; hence a separating graph can indeed serve as a counterexample to derivability.

Lemma 39. *A separating graph for $\Gamma \vdash \mathcal{E}$ w.r.t. $\mathcal{SC}_{\mathcal{L}}$ exists iff $\Gamma \vdash \mathcal{E}$ is not derivable in $\mathcal{SC}_{\mathcal{L}}$.*

Proof. ($\Leftarrow$) If $\Gamma \vdash \mathcal{E}$ is not derivable in $\mathcal{SC}_{\mathcal{L}}$ then there is a failed derivation that ends up in a graph that satisfies the conditions of Def. 38.

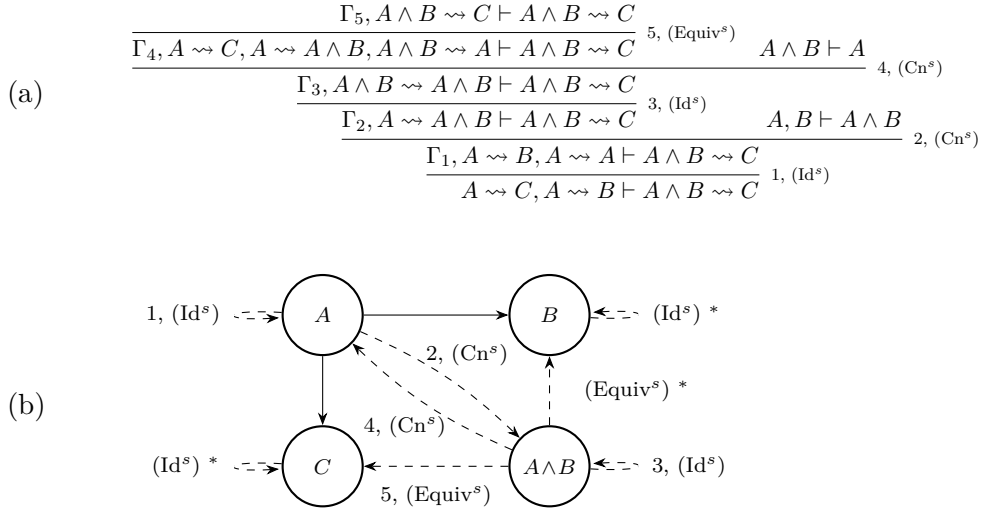


Figure 6.2: Derivation of $A \rightsquigarrow B, A \rightsquigarrow C \vdash A \wedge B \rightsquigarrow C$ in $\mathcal{SC}_C$. (a) Sequent view. (b) Graph view.

($\Rightarrow$) We cannot simultaneously have a separating graph G and a proof in $\mathcal{SC}_C$ for $\Gamma \vdash \mathcal{E}$. Suppose to have both, then there exists a proof branch, whose conditionals on the LHS of every sequent belong to $E(G)$: for the root this is true by condition (2) of Def. 38 and for every rule application there is a premise for which this is true by condition (4). Then this also holds for the sequent-leaf of this branch, which has $\mathcal{E}$ on the LHS, contradicting condition (3) of Def. 38. $\square$

Example 15. Fig. 6.2a depicts the sequent proof of $(A \rightsquigarrow B), (A \rightsquigarrow C) \vdash (A \wedge B \rightsquigarrow C)$ in the calculus $\mathcal{SC}_C$; Γ_i denotes all the conditionals on the LHS apart from the one added at the previous (i -th) step and from those used at the next ($(i+1)$ -th) step. Fig. 6.2b depicts its graph view. The solid arrows represent the initial edges of the graph. The dashed arrows are added by applying the rules, and are labeled with the rule's name. Those corresponding to rule applications in the sequent derivation are also labeled with the step number. $(A \wedge B \rightsquigarrow C)$ is derived at step 5. Not needed to reach a proof, the edges labeled with $*$ are present in the minimal saturated graph.

Consider now the sequent $(A \rightsquigarrow B), (A \rightsquigarrow C) \vdash (C \rightsquigarrow A \wedge B)$, not provable in $\mathcal{SC}_C$. The same graph in Fig. 6.2b corresponds to the failed derivation and it is separating. We can get other separating graphs by introducing extra edges (for example, the edge (B, C) can be added without violating saturation w.r.t. the rules of $\mathcal{SC}_C$).

Henceforth we use both sequent and graph perspectives of our calculi and derivations, depending on which is more convenient at the moment. In the upcoming soundness and completeness proofs, we will mainly employ the sequent view for the former and

the graph view for the latter. We now outline the roadmaps for the soundness and completeness proofs. Subsections for each logic will provide relevant details.

Roadmap of Soundness: For **C**, **CL** and **CM** the proofs proceed by emulating the calculus rules using the KLM principles for the corresponding logic. The rules of $\mathcal{SC}_C$, $\mathcal{SC}_{CL}$, and $\mathcal{SC}_{CM}$ are indeed “sequent-ized” versions of (or derivable from) the original KLM rules. The same does not hold for (CnM^s) , present in **P** and **M**, for which soundness is instead proved semantically.

Roadmap of Completeness: For each KLM logic $\mathcal{L}$, we show that there is no proof of $\Gamma \vdash \mathcal{E}$ in $\mathcal{SC}_{\mathcal{L}}$ only when the given entailment does not hold. As failed derivations end up in a separating graph G (Lem. 39), we use G to define an $\mathcal{L}$ -countermodel $\mathbf{M}_{\mathcal{L}}(G)$ for $\Gamma \vdash \mathcal{E}$. $\mathbf{M}_{\mathcal{L}}(G)$ is constructed to satisfy exactly the conditionals between formulas in $V(G)$ that appear as edges in G .

Definition 39. A preference model M is aligned with a graph G if $M \models A \rightsquigarrow B$ is equivalent to $(A \rightsquigarrow B) \in E(G)$ for all $A, B \in V(G)$.

Lemma 40. If M is aligned with a separating graph G for $\Gamma \vdash \mathcal{E}$ then M is a countermodel for $\Gamma \vdash \mathcal{E}$.

Proof. Since G is a separating graph, $\Gamma \subseteq E(G)$ and $\mathcal{E} \notin E(G)$. Due to alignment, M therefore satisfies all conditionals in Γ , but not $\mathcal{E}$. $\square$

Our countermodel construction is inspired by the canonical model construction in the KLM paper, albeit with a distinct “analytic” approach. We indeed focus exclusively on conditionals between formulas from $V(G)$, rather than all possible conditionals. In particular, the states of our countermodels will be equivalence classes on $V(G)$ (vs on all formulas as in the KLM paper) w.r.t. *conditional equivalence in G* defined as: $A \leftrightarrow B$ when $\{(A, B), (B, A)\} \subseteq E(G)$ (vs $A \sim B$ when $A \vdash B$ and $B \vdash A$ in the KLM paper). Notice that $\leftrightarrow$ is an equivalence relation if a graph G is saturated w.r.t. the rules (Id^s) (for reflexivity) and $(Equiv^s)$ (for transitivity), present in all our calculi.

Henceforth, $[A]_{\leftrightarrow}$ will denote the equivalence class of the formula A and $V(G)/\leftrightarrow$ will denote the set of all equivalence classes.

The labeling function and the preference relation for countermodels will be tailored to each logic $\mathcal{L}$, following a methodology similar to the canonical model constructions in the KLM paper. This will ensure compliance with the conditions imposed on $\mathcal{L}$ -models, relying on saturation w.r.t. the rules of the calculus $\mathcal{SC}_{\mathcal{L}}$ in each case.

The labeling of a state will be constructed to satisfy exactly the consequences of all its formulas in G , as stated in the requirement below. $N_G(A)$ will denote the set $\{B \mid A \rightsquigarrow B \in E(G)\}$, i.e. the set of consequences (or neighbors) in G for all conditionals in $E(G)$. We extend this notation to sets of formulas: $N_G(\alpha) = \bigcup_{A \in \alpha} N_G(A)$, and its complement $\overline{N}_G(\alpha) = V(G) \setminus N_G(\alpha)$.

Requirement 1 (Labeling requirement). *We seek for a labeling function l on $V(G)/\leftrightarrow$ such that for any $\alpha \in V(G)/\leftrightarrow$ and $B \in V(G)$, $l(\alpha) \models B$ iff $B \in N_G(\alpha)$.*

For the monotonic logics **CM** and **M**, the fulfilment of Req. 1 will be enough to prove alignment with G for the degenerate case of an empty preference relation (as required in these logics). For the non-monotonic logics **C**, **CL**, and **P** we chose a preference relation $\preceq$ in such a way that for every A state $[A]_{\leftrightarrow}$ is the least state (and therefore the only minimal state) satisfying A .

Requirement 2 (Preference requirement). *We seek for a model $M = \langle V(G)/\leftrightarrow, \preceq, l \rangle$ such that $[A]_{\leftrightarrow} \in \text{least}_{\preceq}(\|A\|^M)$ for any $A \in V(G)$.*

Satisfying Req. 1 and Req. 2 implies alignment with G , so it is sufficient to establish a countermodel for $\Gamma \vdash \mathcal{E}$.

Lemma 41. *If G is separating graph for $\Gamma \vdash \mathcal{E}$, and $M = \langle V(G)/\leftrightarrow, \preceq, l \rangle$ satisfies Req. 2 and l satisfies Req. 1, then M is a countermodel for $\Gamma \vdash \mathcal{E}$.*

Proof. For $A, B \in V(G)$, $M \models A \rightsquigarrow B$ iff $M, [A]_{\leftrightarrow} \models B$, since $\text{min}_{\preceq}(X) = \{[A]_{\leftrightarrow}\}$ due to Req. 2, which is equivalent to $B \in N_G(A)$ by Req. 1. So M is aligned with G , which implies that M is a countermodel for $\Gamma \vdash \mathcal{E}$ by Lem. 40 $\square$

To achieve completeness, it remains to verify that $\mathbf{M}_{\mathcal{L}}(G)$ is an $\mathcal{L}$ -model. Ensuring smoothness for our countermodels requires special attention. For the KLM logics that mandate the preference relation to be a strict partial order (i.e., all but **C**), smoothness naturally follows from the finiteness of the constructed model (in a finite partially ordered set, every decreasing chain has the least element). The situation for **C** is more intricate. Req. 2 guarantees the existence of the least element in $\|A\|^M$ for all $A \in V(G)$, but not for other formulas. This condition is exactly the *analytic l -smoothness* condition introduced in Sec. 6.4. This alternative version of smoothness (characterization finite entailments in **C**) will therefore be crucial to be able to reconstruct a countermodel from a separating graph.

6.5.1 The base logic C

The calculus $\mathcal{SC}_{\mathbf{C}}$ for **C** consists of the rules (Id^s), ($Equiv^s$), and (Cn^s) from Fig. 6.1. (Id^s) is a “sequent-ized” version of one of the original rules for **C**, (Cn^s) allows the rule (RWk) to be applied to multiple conditionals having the same conditions, while ($Equiv^s$) mirrors the (*Equivalence*) rule in the KLM paper and allows swapping conditionally equivalent conditions. Their soundness follows from the lemma below, stating that for each rule Σ derives the generated conditional δ_1 via the original KLM rules for **C**.

Theorem 42 (Soundness of $\mathcal{SC}_{\mathbf{C}}$). *If $\Gamma \vdash \mathcal{E}$ is derivable in $\mathcal{SC}_{\mathbf{C}}$ then $\Gamma \vdash_{\mathbf{C}} \mathcal{E}$.*

Proof. By induction on the derivation size. For each rule we go from the derivability of the (conditional) premise to the derivability of the conclusion by showing that $\Sigma \vdash_{\mathbf{C}} \delta_1$. For (Id^s) , $\delta_1 = (A, A)$ is derived by (Id) . Rule $\rightarrow$ is emulated as follows:

$$\frac{(A', A) \quad \frac{(A, A') \quad (A, B)}{(A' \wedge A, B)} \text{ (CMon)}}{(A', B)} \text{ (CCut)}$$

For (Cn^s) we use the (And) rule $A \rightsquigarrow B_1, A \rightsquigarrow B_1 \vdash A \rightsquigarrow B_1 \wedge B_2$, derivable in $\mathbf{C}$ as follows:

$$\frac{(A, B_1) \quad (A, B_2) \quad \frac{\overline{(A \wedge B_1 \wedge B_2, A \wedge B_1 \wedge B_2)}}{(A \wedge B_1 \wedge B_2, B_1 \wedge B_2)} \text{ (RWk)}}{(A, B_1) \quad (A \wedge B_1, B_1 \wedge B_2)} \text{ (CCut)} \quad \frac{}{(A, B_1 \wedge B_2)} \text{ (CCut)}$$

The conditional $\delta_1 = (A, C)$ generated by (Cn^s) arises by n applications of (And) followed by (RWk) , since the LK premise implies $C \in Cn(B_1 \wedge \dots \wedge B_n)$. $\square$

For completeness, as described in the roadmap, we build a countermodel $\mathbf{M}_{\mathbf{C}}(G)$ satisfying Req. 1 and Req. 2 for a separating graph G for $\Gamma \vdash \mathcal{E}$ w.r.t. $\mathcal{SC}_{\mathbf{C}}$.

The definition of the labeling function will rely on the saturation of G w.r.t. (Cn^s) . Hence $N_G(A)$ contains all the formulas from $V(G)$ that are classically entailed by $N_G(A)$. So, if $B \in \overline{N}_G(A)$ then $B \notin Cn(N_G(A))$. We can lift it to the level of equivalence classes, since $N_G([A]_{\leftrightarrow}) = N_G(A)$ due to the saturation w.r.t. $(Equiv^s)$: $B \notin Cn(N_G(\alpha))$ for any $\alpha \in V(G)/\leftrightarrow$ and $B \in \overline{N}_G(\alpha)$. Hence there exist Boolean assignments satisfying all formulas in $N_G(\alpha)$ and falsifying B . We denote by $\rho(\alpha, B)$ one (arbitrarily chosen) such assignment. We can then define a labeling function $W_G(\alpha) = \{\rho(\alpha, B) \mid B \in \overline{N}_G(\alpha)\}$, satisfying Req. 1 and mapping each state into linearly many Boolean assignments.

We also reuse the definition of preference relation from the canonical model for $\mathbf{C}$ in the KLM paper, adapting it to our analytic setting.

Definition 40. For $\alpha, \beta \in V(G)/\leftrightarrow$, $\beta \preceq_G \alpha$ when there exists $B \in \beta$ such that $B \in N_G(\alpha)$.

Notice that $\preceq_G$ is antisymmetric for G saturated w.r.t. $(Equiv^s)$. Indeed $\alpha \preceq_G \beta$ and $\beta \preceq_G \alpha$ would imply $A \leftrightarrow B$ for some $A \in \alpha$ and $B \in \beta$, implying $\alpha = \beta$.

Countermodel for logic $\mathbf{C}$ is then defined as follows:

$$\mathbf{M}_{\mathbf{C}}(G) = \langle V(G)/\leftrightarrow, \preceq_G, W_G \rangle$$

Lemma 42. If G is a separating graph for $\Gamma \vdash \mathcal{E}$ w.r.t. $\mathcal{SC}_{\mathbf{C}}$ then $\mathbf{M}_{\mathbf{C}}(G)$ is a countermodel for $\Gamma \vdash \mathcal{E}$.

Proof. W_G satisfies Req. 1 by definition. $\mathbf{M}_C(G)$ satisfies Req. 2 since for every $B \in V(G)$ and for every $\alpha \in V(G)/\leftrightarrow$ such that $\alpha \neq [B]_{\leftrightarrow}$ if $\alpha \in \|B\|^{\mathbf{M}_C(G)}$ then $B \in N_G(\alpha)$ due to Req. 1, so $[B]_{\leftrightarrow} \prec_G \alpha$ by definition and antisymmetry of $\preceq_G$, so $[B]_{\leftrightarrow} \in \text{least}_{\prec}(\|B\|^{\mathbf{M}_C(G)})$. The proof is concluded by Lem. 41. $\square$

$\mathbf{M}_C(G)$ satisfies analytic l-smoothness, because of Req. 2, therefore we can apply Th. 41 for completeness.

Theorem 43 (Completeness of $\mathcal{SC}_C$). *If $\Gamma \vdash_C \mathcal{E}$ then $\Gamma \vdash \mathcal{E}$ is derivable in $\mathcal{SC}_C$.*

Proof. By contraposition: if $\Gamma \vdash \mathcal{E}$ is not derivable in $\mathcal{SC}_C$, by Lem. 39 there is a separating graph G for it. Hence we can construct a countermodel $\mathbf{M}_C(G)$ satisfying analytic l-smoothness for $\Gamma \cup \{\mathcal{E}\}$ such that $\mathbf{M}_C(G) \not\models \mathcal{E}$ and $\mathbf{M}_C(G) \models \mathcal{C}$ for every $\mathcal{C} \in \Gamma$. By Th. 41 it follows that $\Gamma \not\vdash_C \mathcal{E}$. $\square$

Example 16. *Returning to Ex. 15, consider the separating graph G from Fig. 6.2 for*

$$A \rightsquigarrow B, A \rightsquigarrow C \vdash C \rightsquigarrow A \wedge B.$$

Its equivalence classes are

$$[A]_{\leftrightarrow} = [A \wedge B]_{\leftrightarrow} = \{A, A \wedge B\}, \quad [B]_{\leftrightarrow} = \{B\}, \quad [C]_{\leftrightarrow} = \{C\}.$$

Moreover,

$$N_G([A]_{\leftrightarrow}) = \{A, B, C, A \wedge B\}, \quad N_G([B]_{\leftrightarrow}) = \{B\}, \quad N_G([C]_{\leftrightarrow}) = \{C\}.$$

Hence the reconstructed model $M_C(G)$ has three states, with $[B]_{\leftrightarrow} \prec_G [A]_{\leftrightarrow}$, $[C]_{\leftrightarrow} \prec_G [A]_{\leftrightarrow}$, while $[B]_{\leftrightarrow}$ and $[C]_{\leftrightarrow}$ are incomparable.

By the definition of W_G , each state satisfies exactly its G -neighbours among the formulas occurring in the graph. Thus $[A]_{\leftrightarrow}$ is the unique minimal A -state and satisfies both B and C , so $M_C(G) \models A \rightsquigarrow B$ and $M_C(G) \models A \rightsquigarrow C$. On the other hand, $[C]_{\leftrightarrow}$ is the unique minimal C -state and does not satisfy $A \wedge B$. Therefore $M_C(G) \not\models C \rightsquigarrow A \wedge B$, as required.

6.5.2 Logics CL and CM

The calculi $\mathcal{SC}_{CL}$ and $\mathcal{SC}_{CM}$ for the **CL** and **CM** extend $\mathcal{SC}_C$ with the rule ($Loop^s$) and the transitivity rule (Tr^s), respectively. Note that ($Loop^s$) is derivable from (Tr^s).

Theorem 44 (Soundness of $\mathcal{SC}_{CL}$ and $\mathcal{SC}_{CM}$). *If $\Gamma \vdash \mathcal{E}$ is derivable in $\mathcal{SC}_{CL}$ (resp. $\mathcal{SC}_{CM}$) then $\Gamma \vdash_{CL} \mathcal{E}$ ($\Gamma \vdash_{CM} \mathcal{E}$).*

Proof. Similar to the proof of Th. 42. Observe that $(Loop^s)$ can be emulated in **CL** directly by $(Loop)$ principle, and $(Transitivity)$ can be emulated in **CM** as follows:

$$\frac{(A, B) \quad \frac{(B, C)}{(A \wedge B, C)}^{(Mon)}}{(A, C)}^{(CCut)}$$

□

For the countermodels $\mathbf{M}_{\mathbf{CL}}(G)$ and $\mathbf{M}_{\mathbf{CM}}(G)$ we re-define the preference relation in the countermodel $\mathbf{M}_{\mathbf{C}}(G)$ to fulfill the semantic requirements for these logics – transitivity for **CL** and the empty relation for **CM** – using the saturation w.r.t. the additional rules $(Loop^s)$ and (Tr^s) .

Following the KLM paper, transitivity for **CL** is ensured by taking the (positive) transitive closure of $\prec_G$ as the preference relation.

Definition 41. *Positive transitive closure of a relation $\prec_G$ is defined as follows: $x \prec_G^+ y$ when there is a chain of worlds $w_0 \prec_G \dots \prec_G w_n$ such that $w_0 = x$, $w_n = y$, and $n \geq 1$.*

The countermodel for **CL** is therefore defined as follows:

$$\mathbf{M}_{\mathbf{CL}}(G) = \langle V(G)/\leftrightarrow, \prec_G^+, W_G \rangle$$

Saturation w.r.t. $(Loop^s)$ ensures that the transitive closure is asymmetric.

Lemma 43. *If G is saturated w.r.t. $(Loop^s)$ and $(Equiv^s)$, $\prec_G^+$ is asymmetric.*

Proof. Suppose $\alpha \prec_G^+ \beta$ and $\beta \prec_G^+ \alpha$ for some $\alpha \neq \beta$. By the definition of $\prec_G$ and saturation w.r.t. $(Equiv^s)$ there exists a directed cycle in G going through some $A \in \alpha$ and some $B \in \beta$. Due to the saturation w.r.t. $(Loop^s)$, for each edge of the cycle its inverse edge also belongs to $E(G)$, then due to saturation w.r.t. $(Equiv^s)$ we have $A \leftrightarrow B$, so $\alpha = \beta$, which contradicts the initial assumption. □

Since $\mathbf{M}_{\mathbf{CL}}(G)$ has the same labeling as $\mathbf{M}_{\mathbf{C}}(G)$ and Req. 2 is preserved by taking a positive transitive closure (it preserves the least elements) we have completeness.

Theorem 45 (Completeness of $\mathcal{SC}_{\mathbf{CL}}$). *If $\Gamma \vdash_{\mathbf{CL}} \mathcal{E}$ then $\Gamma \vdash \mathcal{E}$ is derivable in $\mathcal{SC}_{\mathbf{CL}}$.*

Proof. If $\Gamma \vdash \mathcal{E}$ is not derivable, a separating graph G exists by Lem. 39. $\mathbf{M}_{\mathbf{CL}}(G)$ satisfies Req. 1 and 2, so it is a countermodel for $\Gamma \vdash_{\mathbf{CM}} \mathcal{E}$ by Lem. 41. $\prec_G^+$, being asymmetric and transitive (a strict partial order), ensures smoothness due to the finiteness of $\mathbf{M}_{\mathbf{CL}}(G)$, so $\mathbf{M}_{\mathbf{CL}}(G)$ is a **CL**-countermodel for $\Gamma \vdash \mathcal{E}$. □

For **CM**, we take an empty preference relation $\preceq_\emptyset$:

$$\mathbf{M}_{\mathbf{CM}}(G) = \langle V(G)/\leftrightarrow, \preceq_\emptyset, W_G \rangle$$

Req. 1, satisfied by W_G , and saturation w.r.t. (Tr^s) are enough to prove alignment with G (and, therefore, completeness) directly.

Theorem 46 (Completeness of $\mathcal{SC}_{\mathbf{CM}}$). *If $\Gamma \vdash_{\mathbf{CM}} \mathcal{E}$ then $\Gamma \vdash \mathcal{E}$ is derivable in $\mathcal{SC}_{\mathbf{CM}}$.*

Proof. If $\Gamma \vdash \mathcal{E}$ is not derivable, by Lem. 39 there is a separating graph G . Since $\mathbf{M}_{\mathbf{CM}}(G)$ has an empty preference relation, it is a **CM**-model and all states within each truth set are minimal. By Lem. 40 it suffices to show that $\mathbf{M}_{\mathbf{CM}}(G)$ is aligned with G . If $(A' \rightsquigarrow B') \notin E(G)$ then $\mathbf{M}_{\mathbf{CM}}(G), [A']_{\leftrightarrow} \models A'$ and $\mathbf{M}_{\mathbf{CM}}(G), [A']_{\leftrightarrow} \not\models B'$ (both due to Req. 1 satisfied by W_G), so $\mathbf{M}_{\mathbf{CM}}(G) \not\models A' \rightsquigarrow B'$. Now let us show that $(A \rightsquigarrow B) \in E(G)$ implies $\mathbf{M}_{\mathbf{CM}}(G) \models A \rightsquigarrow B$. For any $\alpha \in \min_{\prec}(\|A\|^{\mathbf{M}_{\mathbf{CM}}(G)})$ we have $A \in N_G(\alpha)$ due to Req. 1 satisfied by W_G , i.e. there is $C \in \alpha$ such that $(C \rightsquigarrow A) \in E(G)$. Due to saturation of G w.r.t. (Tr^s) the edges of G are transitive, so $(C \rightsquigarrow B) \in E(G)$, which implies $\mathbf{M}_{\mathbf{CM}}(G), \alpha \models B$ again by Req. 1. $\square$

6.5.3 Logics **P** and **M**

The calculi $\mathcal{SC}_{\mathbf{P}}$ and $\mathcal{SC}_{\mathbf{M}}$ for **P** and **M**¹ are obtained by extending $\mathcal{SC}_{\mathbf{CL}}$ and $\mathcal{SC}_{\mathbf{CM}}$, with the (CnM^s) rule with $m + 1$ premises:

$$\frac{\{\Pi, A \rightsquigarrow B_1, \dots, A \rightsquigarrow B_n, A \rightsquigarrow C_j \vdash \mathcal{E}\}_{j=1}^m \quad B_1, \dots, B_n \vdash C_1, \dots, C_m}{\Pi, A \rightsquigarrow B_1, \dots, A \rightsquigarrow B_n \vdash \mathcal{E}}$$

Note that (Cn^s) is an instance of (CnM^s) for $m = 1$. The motivation for (CnM^s) stems from a semantic perspective. Indeed **P** and **M** differ from the cumulative logics **CL** and **CM** by the restriction that each state is labeled with exactly one Boolean assignment. Hence, to extend our approach to **P** and **M** we need to replace the set $W_G(\alpha)$ of assignments falsifying each formula from α individually with a single assignment falsifying all of them simultaneously. This is achieved by transforming the rule (Cn^s) to a multi-premise version (CnM^s) . Saturation w.r.t. this rule ensures the existence of one assignment satisfying Req. 1: take the instance of the rule (CnM^s) instantiating Σ with $\{A \rightsquigarrow B \mid B \in N_G(\alpha)\}$ and $\{\delta_i\}_{i=1}^m$ with $\{A \rightsquigarrow C \mid C \in \overline{N}_G(\alpha)\}$, then the side condition $\mathfrak{S}$ is $N_G(\alpha) \vdash \overline{N}_G(\alpha)$ and it should be false due to saturation w.r.t. (CnM^s) . Therefore there exists a Boolean assignment satisfying all formulas in $N_G(\alpha)$ and falsifying all formulas in $\overline{N}_G(\alpha)$ (thus satisfying Req. 1). We will denote such assignment $W_G^1(\alpha)$.

We now replace W_G with W_G^1 in the definition of countermodel for **CL** and **CM** and get a countermodel in which each state is labeled by exactly one Boolean assignment:

¹We include **M** (= classical logic) in our investigation for sake of completeness, and to highlight the uniformity of our approach.

- $\mathbf{M_P}(G) = \langle V(G)/\leftrightarrow, \prec_G^+, W_G^1 \rangle$.
- $\mathbf{M_M}(G) = \langle V(G)/\leftrightarrow, \prec_\emptyset, W_G^1 \rangle$.

Since we changed only labeling, by preserving Req. 1 and simultaneously ensuring functionality, the completeness proofs in Th. 45 and Th. 46 easily extend to **P** and **M**.

Theorem 47 (Completeness). *If $\Gamma \vdash_{\mathbf{P}} \mathcal{E}$ (resp. $\Gamma \vdash_{\mathbf{M}} \mathcal{E}$) then $\Gamma \vdash \mathcal{E}$ is derivable in $\mathcal{SC_P}$ (resp. $\mathcal{SC_M}$).*

The soundness of the rule (CnM^s) for **P** and **M** is proved semantically using the characterization of finite entailments in terms of l-smoothness (see Remark 12,p.140).

Theorem 48 (Soundness of $\mathcal{SC_P}$ and $\mathcal{SC_M}$). *If $\Gamma \vdash \mathcal{E}$ is derivable in $\mathcal{SC_L}$ then $\Gamma \vdash_{\mathcal{L}} \mathcal{E}$, for $\mathcal{L} \in \{\mathbf{P}, \mathbf{M}\}$.*

Proof. By induction on the length of the proof. If the last applied rule is a concluding axiom, the claim is trivial; for the rules other than (CnM^s) it follows from the demonstrated soundness w.r.t. the weaker KLM logics. The proof for (CnM^s) is by contradiction. If $\Gamma \not\vdash_{\mathcal{L}} \mathcal{E}$ then there is a $\mathcal{L}$ -countermodel M for $\Gamma \vdash \mathcal{E}$ satisfying l-smoothness. For the rule (CnM^s) , we assume $\{A \rightsquigarrow B_i\}_{i=1}^n \subseteq \Gamma$, and $\bigvee_{j=1}^m C_j \in \text{Cn}(\{B_i\}_{i=1}^n)$. In M there is exactly one state in $\text{least}_{\prec}(\|A\|^M)$ labeled with one world, which satisfies B_i for every i , so it should satisfy also $\bigvee_{j=1}^m C_j$, and therefore C_k for some k . Then M satisfies $A \rightsquigarrow C_k$ and thus is a $\mathcal{L}$ -countermodel for a premise $\Gamma, A \rightsquigarrow C_k \vdash \mathcal{E}$, which contradicts (one of) the inductive hypotheses. $\square$

6.6 Complexity: all KLM logics are co-NP-complete

We utilize the introduced calculi to establish uniform complexity results for the entailment problem in KLM logics, as well as for two useful restrictions on the form of conditionals: Horn and literal conditionals. Remember that in Horn (resp. literal) conditional $A \rightsquigarrow B$, A is a conjunction of propositional variables (resp. literals, i.e. possibly negated variables) and B is a single propositional variable (resp. literal).

As seen before, entailment checking in a KLM logic $\mathcal{L}$ can be reduced to the search for a separating graph G w.r.t. $\mathcal{SC_L}$ (Lem. 39). Note that the saturation of G w.r.t. every rule can be checked in polynomial time *modulo checks of classical entailment conditions* $\mathfrak{S}$.

Lemma 44. *Checking the saturation of graph w.r.t. $\mathcal{SC_L}$ can be reduced in polynomial time to a polynomial number of checks of classical non-entailment.*

Proof. Saturation w.r.t. (Id^s) , (Equiv^s) and (Tr^s) can be checked directly by examining all (triplets of) nodes in $V(G)$. Saturation w.r.t. (Loop^s) means that there is no cycle in G going through several equivalence classes of $V(G)/\leftrightarrow$, which can be easily checked in

polynomial time. For (Cn^s) and (CnM^s) we check violations of the side condition $\mathfrak{S}$ for subset-maximal choices of Σ and $\{\delta_j\}_{j=1}^m$, since removing conditionals from Σ or $\{\delta_j\}_{j=1}^m$ cannot make $\mathfrak{S}$ valid. Hence for (Cn^s) we check $B \notin Cn(N_G(A))$ for all $A \in V(G)$ and $B \in \overline{N}_G(A)$, and for (CnM^s) we check $\bigvee_{B \in \overline{N}_G(A)} B \notin Cn(N_G(A))$ for all $A \in V(G)$. $\square$

Thus, we get co-NP-completeness with a separating graph playing the role of a witness for non-entailment.

Theorem 49. *For each $\mathcal{L} \in \{\mathbf{C}, \mathbf{CL}, \mathbf{CM}, \mathbf{P}, \mathbf{M}\}$, the entailment problem in $\mathcal{L}$ is co-NP-complete.*

Proof. $\Gamma \not\vdash_{\mathcal{L}} \mathcal{E}$ can be checked in non-deterministic polynomial time by guessing a separating graph G for $\Gamma \vdash \mathcal{E}$ and checking its saturation w.r.t. $\mathcal{SC}_{\mathcal{L}}$ as in Lem. 44 and guessing falsifying assignments for the resulting classical entailments. Co-NP-hardness is given by the following reduction: A is classically valid iff $\vdash_{\mathcal{L}} \top \rightsquigarrow A$. $\square$

This yields tractable fragments when the shape of conditionals is constrained so that the classical entailments in S can be verified in polynomial time.

Theorem 50. *For $\mathcal{L} \in \{\mathbf{C}, \mathbf{CL}, \mathbf{CM}\}$, the entailment problem for Horn and literal conditionals in $\mathcal{L}$ is in PTIME.*

Proof. Entailment between conjunctions of literals is verifiable in polynomial time. Hence so is the check of possible rule applications using Lem. 44. For $\mathcal{L} \in \{\mathbf{C}, \mathbf{CL}, \mathbf{CM}\}$, derivations in $\mathcal{SC}_{\mathcal{L}}$ consist of a single branch, so either a proof or a separating graph is found in polynomial time. $\square$

As shown in Th. 52, this result extends to $\mathbf{P}$ (and $\mathbf{M}$) for Horn conditionals, although it does not apply to literal conditionals.

Theorem 51. *For $\mathcal{L} \in \{\mathbf{P}, \mathbf{M}\}$, the entailment problem in $\mathcal{L}$ for literal conditionals is co-NP-hard.*

Proof. By a reduction of CNF-SAT to non-entailment in $\mathcal{L}$. An arbitrary CNF-formula $F = \bigwedge_{i=1}^m \bigvee_{j=1}^{n_i} l_{ij}$ is classically satisfiable iff $\Gamma(F) \not\vdash_{\mathcal{L}} \top \rightsquigarrow \perp$ where

$$\Gamma(F) = \{(\bigwedge_{j=1}^{n_i} \neg l_{ij}) \rightsquigarrow \perp\}_{i=1}^m$$

(with double negations simplified). The satisfying assignment for F corresponds to the labeling of $\top$ -minimal state in the countermodel for KLM-entailment, which has to satisfy at least one l_{ij} for every i . $\square$

Hence for KLM logics, the inclusion of (Or) results in increased complexity for entailment of literal conditionals.

The KLM paper semantically shows that **P** and **CL** derive identical conditionals in the Horn fragment. This and the analogous result for **M** and **CM**, can be seen using our calculi too.

Theorem 52. *If $\Gamma \cup \{\mathcal{E}\}$ is finite and contains only Horn conditionals, then $\Gamma \vdash_{\mathbf{CL}} \mathcal{E}$ (resp. $\Gamma \vdash_{\mathbf{CM}} \mathcal{E}$) iff $\Gamma \vdash_{\mathbf{P}} \mathcal{E}$ (resp. $\Gamma \vdash_{\mathbf{M}} \mathcal{E}$).*

Proof. If $\{B_i\}_{i=1}^n$, and $\{C_j\}_{j=1}^m$ are conjunctions of variables, $B_1, \dots, B_n \vdash C_1, \dots, C_m$ is derivable in **LK** iff some C_k contains only variables occurring on the antecedent, which implies derivability of $B_1, \dots, B_n \vdash C_k$. So for Horn conditionals we can replace each application of (CnM^s) with an application of (Cn^s) (removing all sequent premises but one), obtaining a derivation in the corresponding calculus without (Or). $\square$

6.7 Prover: KLMPPropagator

The saturated calculi developed in this chapter were implemented in the automated theorem prover KLMPPropagator², developed by Clemens Eisenhofer. Rather than performing proof search directly in the sequent calculi, the prover exploits the graph view of them. An entailment problem is translated into the search for a separating graph, whose existence is equivalent to non-derivability in the corresponding calculus (Lem. 39). The graph is represented by propositional variables encoding conditional dependencies between formulas, while the saturation conditions induced by the calculus rules are implemented as SMT propagations. As a result, determining entailment is reduced to checking the satisfiability of an SMT encoding of the existence of a separating graph. If no such graph exists, the entailment is proved; otherwise, the resulting graph can be used to reconstruct a countermodel.

The implementation uses the Z3 SMT solver and its user-propagator framework. The encoding directly reflects the saturating calculi: graph edges correspond to conditional formulas, while rule applications are realized as propagations enforcing saturation of the graph. This design allows the prover to generate countermodels for non-entailments and naturally extends to richer underlying theories supported by SMT technology.

Experimental evaluation in [CER24], given in Fig. 6.3, compared KLMPPropagator with several existing provers for KLM logics, including KLMLearn 2.0, NESCOND, and IOCondProver. On randomly generated benchmark instances, KLMPPropagator solved almost all test cases within the imposed ten-second limit, whereas the competing systems frequently timed out, especially on larger instances. The results indicate that the SMT-based approach significantly outperforms previous tableau- and sequent-based implementations and demonstrate the practical effectiveness of the saturating graph characterizations introduced in this work.

²<https://github.com/CEisenhofer/KLMPPropagator>

5 conditional premises				
Logic	<i>KLMPpropagator</i>	KLMlean 2.0	NESCOND	IOCondProver
C	100 (82 + 18)	3 (0 + 3)	22 (7 + 15)	-
CL	100 (82 + 18)	23 (5 + 18)	-	-
P	100 (75 + 25)	41 (16 + 25)	-	-
CM	100 (76 + 24)	-	-	98 (74 + 24)
20 conditional premises				
Logic	<i>KLMPpropagator</i>	KLMlean 2.0	NESCOND	IOCondProver
C	100 (47 + 53)	4 (0 + 4)	18 (0 + 18)	-
CL	98 (45 + 53)	28 (0 + 28)	-	-
P	100 (13 + 87)	50 (0 + 50)	-	-
CM	100 (23 + 77)	-	-	33 (10 + 23)

Figure 6.3: Experimental evaluation for KLM provers from [CER24]. Number of instances of randomly generated entailment problems in different KLM logics (out of 100) solved by different provers within 10s (SAT + UNSAT)

The development of KLMPpropagator also provides evidence that the proof-theoretic characterization we developed for KLM logics is not only useful for establishing theoretical properties, but can also serve as a foundation for efficient automated reasoning systems.

Globally non-monotonic reasoning

This chapter turns to *globally* non-monotonic reasoning, focusing on the influential approach called *rational closure*. While the KLM logics studied in the previous chapter successfully isolate and characterize local non-monotonicity, when viewed as operators of conditional entailment they remain globally monotonic and therefore inadequate for many practical reasoning tasks. In particular, they do not support revision of conclusions in light of new information: once a conditional follows from a knowledge base, it continues to follow from every extension of that knowledge base. To address this limitation, Lehmann and Magidor [LM92] proposed rational closure, a globally non-monotonic entailment operator built on top of the KLM framework, as a canonical solution to the problem of entailment from conditional knowledge bases. Rational closure has since established itself as one of the most influential formalisms in non-monotonic reasoning, shaping both the theoretical study of conditional entailment and its applications in knowledge representation. Yet, despite its importance and well-understood semantic and constructive characterizations, it has received surprisingly little attention from a proof-theoretic perspective.

In this chapter, we extend the analytic methodology developed in the previous chapters from locally to globally non-monotonic reasoning. Specifically, we apply the **GL**-based approach of [GGOP09] employed in Chapters 3 and 4, to obtain the first analytic proof calculus for rational closure. The chapter begins with a discussion of the origins of rational closure and its significance (Sec. 7.1) and the overview of various existing characterizations of it (Sec. 7.2). We then formally introduce rational closure from two principal complementary perspectives: a constructive one and semantic one (Sec. 7.3). As a first step of our analytic development, we enrich the **GL**-based calculus for **P** from [GGOP09] with hypersequents and analytic cuts, encoding comprehensive information about the countermodels in saturated branches of derivations (Section 7.4). This enrichment allows us to extend the calculus in Section 7.5 with an additional rule enforcing model minimality to capture rational closure. The resulting calculus thus provides a modular

proof-theoretic account of both preferential reasoning in KLM logic **P** and entailment by rational closure, with the transition between the two governed by a single rule. As part of the completeness proof, we establish a three-way correspondence between proof-theoretic, constructive, and semantic perspectives on rational closure.

7.1 Background

While the KLM hierarchy laid the conceptual foundations of non-monotonic inference, one of the central subsequent developments in this area turned to a more practical problem of entailment from conditional knowledge bases. Given a set of conditionals representing incomplete knowledge about the world, the question is which further conditionals should be considered entailed, in a way that appropriately handles exceptions and remains open to revision by new information. KLM logics themselves, when treated as entailment operators on conditionals, are not well-suited for this task. First, they capture only a local form of non-monotonicity — the failure of strengthening of the antecedent — while remaining globally monotonic: the entailment of a conditional cannot be violated by extending the knowledge base. Second, local non-monotonicity they exhibit is too conservative for practical reasoning. For instance, a knowledge base containing only the conditional $\text{bird} \rightsquigarrow \text{flies}$ does not entail $(\text{bird} \wedge \text{green}) \rightsquigarrow \text{flies}$, even though nothing in the knowledge base suggests that greenness constitutes an exception.

To address this, Lehmann and Magidor [LM92] proposed rational monotonicity as an additional principle governing entailment from conditional knowledge bases: if $(A \rightsquigarrow C)$ is entailed and $(A \rightsquigarrow \neg B)$ is not, then $(A \wedge B \rightsquigarrow C)$ should be entailed as well. They introduced the rational closure operator — the canonical construction for conditional entailment satisfying this principle.

Rational closure has since become a central reference point in the study of conditional entailment in several respects. The applicability of rational monotonicity itself has been widely debated: Stalnaker [Sta94] presented examples suggesting that it may be too strong in some reasoning scenarios, while Makinson [Mak94] proposed the weaker principle of *Disjunctive Rationality*; more recently, Booth and Varzinczak [BV21] developed the corresponding notion of *disjunctive rational closure*. At the same time, rational closure has served as a starting point for further closure constructions and refinements, including lexicographic closure [Leh99] and multipreference closure [GG21], and has been extensively adapted to default reasoning in more expressive formalisms, in particular description logics [CS10, GOP13, Bon19, BCM⁺21].

7.2 State of the art

Work on rational closure has developed along two complementary perspectives: a constructive one, based on the iterative computation of exceptionality ranks [LM92], and a semantic one, based on the selection of minimal models within the subclass of preferential models known as ranked models [GGOP15]. The constructive perspective in particular

yields an effective decision procedure for entailment with respect to rational closure, matching the exact P^{NP} complexity established in [EL00].

The proof-theoretic perspective, by contrast, remains underexplored. Note that simply adding rational monotonicity as an axiom to the KLM principles does not yield a complete axiomatisation for rational closure: the resulting logic — known as **R** — in fact coincides with KLM logic **P** w.r.t. entailment of conditionals from conditional knowledge bases, diverging only when Boolean combinations of conditionals are considered [GGOP09]. To our knowledge, the only proof calculus for rational closure is that of Straßer [Str09], developed within the highly expressive framework of adaptive logic. There, global non-monotonicity is handled via dynamic derivations in which conclusions may be retracted in light of subsequent reasoning steps, resulting in a proof system that is both non-analytic and quite complex.

7.3 Preliminaries: Two perspectives on rational closure

A conditional knowledge base is a set of conditionals of the form $A \rightsquigarrow B$, where A and B are Boolean formulas. For our proof-theoretic purposes, we consider only finite conditional knowledge bases (we will refer to them simply as *bases*) and study two conditional entailment operators that map any base $\mathcal{K}$ to the set of entailed conditionals: *preferential entailment* (corresponding to KLM preferential logic **P** from Chapter 6) and *rational closure*. In this section, we introduce them from two well-known complementary perspectives: a constructive one and a semantic one.

7.3.1 Constructive perspective

Lehmann and Magidor [LM92] introduced rational closure as the minimal extension of a base satisfying KLM postulates and the following rational monotonicity principle:

(RMon) if $(A \rightsquigarrow C)$ is derivable and $(A \rightsquigarrow \neg B)$ is not, then $(A \wedge B \rightsquigarrow C)$ is derivable

They showed that such an extension exists for a large class of conditional bases (including all finite ones) and provided an explicit construction based on the following notion of *exceptionality ranks*.

Definition 42 (Exceptionality ranks [LM92]). *Formula A is called exceptional for a base $\mathcal{K}$ when $\overline{\mathcal{K}}$ classically entails $\neg A$, where $\overline{\mathcal{K}}$ is the materialization of $\mathcal{K}$ defined as*

$$\overline{\mathcal{K}} = \{A \rightarrow B \mid (A \rightsquigarrow B) \in \mathcal{K}\}.$$

The stratification $RSeq(\mathcal{K}_0)$ of $\mathcal{K}_0$ w.r.t. exceptionality ranks is a sequence $\{\mathcal{C}_i\}_{i=0}^{\infty}$ where $\mathcal{C}_0 = \mathcal{K}_0$ and $\mathcal{C}_{i+1} = E(\mathcal{C}_i)$ for the following notion $E(\mathcal{K})$ of exceptional sub-base:

$$E(\mathcal{K}) = \{(A \rightsquigarrow B) \in \mathcal{K} \mid A \text{ is exceptional for } \mathcal{K}\}.$$

The exceptionality rank of an arbitrary formula A is defined as

$$\mathcal{R}_{\mathcal{K}}^{ex}(A) = \min\{i \mid A \text{ is not exceptional for } \mathcal{C}_i\},$$

with $\mathcal{R}_{\mathcal{K}}^{ex}(A) = \infty$ if A is exceptional for every $\mathcal{C}_i$.

The rational closure can be then defined as the set of conditionals $A \rightsquigarrow B$ for which A is less exceptional than $A \wedge \neg B$.

Definition 43 (Rational closure [LM92]). *Rational closure of a base $\mathcal{K}$ is defined as*

$$RC(\mathcal{K}) = \{X \rightsquigarrow Y \mid \mathcal{R}_{\mathcal{K}}^{ex}(X) < \mathcal{R}_{\mathcal{K}}^{ex}(X \wedge \neg Y)\},$$

assuming the convention $\infty < \infty$.

Example 17. We will use the classical “penguin triangle” from [LM92] as the running example. Consider the base

$$\mathcal{K}_{\Delta} = \{\mathbf{b} \rightsquigarrow \mathbf{f}, \mathbf{p} \rightsquigarrow \neg \mathbf{f}, \mathbf{p} \rightsquigarrow \mathbf{b}\},$$

encoding the statements “birds normally fly”, “penguins normally do not fly”, and “penguins normally are birds”. $\mathbf{p}$ is not consistent with $\overline{\mathcal{K}_{\Delta}}$, while $\mathbf{b}$ is, so $\mathcal{K}_{\Delta}$ is stratified as $\mathcal{C}_0 = \mathcal{K}_{\Delta}$, $\mathcal{C}_1 = \{\mathbf{p} \rightsquigarrow \neg \mathbf{f}, \mathbf{p} \rightsquigarrow \mathbf{b}\}$, $\mathcal{C}_2 = \emptyset$. Consequently, the following conditionals belong to rational closure of $\mathcal{K}_{\Delta}$: $(\mathbf{b} \rightsquigarrow \neg \mathbf{p})$ (“bird is normally not a penguin”), since

$$0 = \mathcal{R}_{\mathcal{K}_{\Delta}}^{ex}(\mathbf{b}) < \mathcal{R}_{\mathcal{K}_{\Delta}}^{ex}(\mathbf{b} \wedge \mathbf{p}) = 1,$$

and $(\mathbf{b} \wedge \mathbf{g} \rightsquigarrow \mathbf{f})$ (“green birds normally fly”), since

$$0 = \mathcal{R}_{\mathcal{K}_{\Delta}}^{ex}(\mathbf{b} \wedge \mathbf{g}) < \mathcal{R}_{\mathcal{K}_{\Delta}}^{ex}(\mathbf{b} \wedge \mathbf{g} \wedge \neg \mathbf{f}) = 1.$$

Def. 43 also provides an efficient decision procedure for rational closure: since $E(\mathcal{K}) \subseteq \mathcal{K}$ in Def. 42, the sequence $RSeq(\mathcal{K})$ reaches a fixed point $\mathcal{C}_{\infty} = \bigcap_{i \in \mathbb{N}} \mathcal{C}_i$ for any finite $\mathcal{K}$ in at most $|\mathcal{K}|$ steps. Therefore, the prefix of $RSeq(\mathcal{K})$ up to the fixed point can be computed directly by definition via a polynomial number of classical entailment checks. The computation of $\mathcal{R}_{\mathcal{K}}^{ex}(A)$ can then also be done by a polynomial number of classical entailment checks, resulting in a procedure deciding membership in $RC(\mathcal{K})$ with complexity P^{NP} . It was later shown that membership in rational closure is actually P^{NP} -complete [EL00].

7.3.2 Semantic perspective

The semantics of rational closure is based on *ranked models*, a special case of preference models with total preference relation. We use the definition from [LM92], equivalently representing modular orders as strict versions of total preorders (as usual in conditional logics).

Definition 44. A ranked model is a tuple $\langle W, \preceq, V \rangle$, where W is a set of worlds, $\preceq$ is a reflexive, transitive, total, and smooth (Def. 34, p.138) relation on W , and $V: W \rightarrow 2^{Atm}$ is a valuation function. $\prec$, the strict version of $\preceq$, is defined by $w_1 \prec w_2$ iff $w_1 \preceq w_2$ and $w_2 \not\preceq w_1$. A ranked model $\langle W, \preceq, V \rangle$ satisfies a conditional $A \rightsquigarrow B$ when $V(u) \models B$ for every $u \in W$ such that $V(u) \models A$ and there is no $w \prec u$ such that $V(w) \models A$. A ranked model M is called a model of a base $\mathcal{K}$ (denoted $M \models \mathcal{K}$) when it satisfies all conditionals in $\mathcal{K}$.

A ranked model can equivalently be defined by ranks of worlds instead of the relation $\preceq$: $\mathcal{R}^M(w_0)$ is defined as the maximal length of a descending chain of worlds $w_0 \succ w_1 \succ \dots$ (starting with rank 0 for minimal worlds).¹ We can then also define the rank of a valuation $\rho \in 2^{Atm}$ as $\mathcal{R}^M(\rho) = \min\{\mathcal{R}^M(w) \mid V(w) = \rho\}$, and the rank of a formula as $\mathcal{R}^M(A) = \min\{\mathcal{R}^M(w) \mid M, w \models A\}$.

Preferential entailment is entailment of conditionals in KLM logic **P**. Although preference relation in models of **P** is not necessarily total, it is shown in [LM92] that the entailment w.r.t. the class of ranked models coincides with preferential entailment. Accordingly, in this chapter, we will use this semantic definition of preferential entailment to facilitate modular treatment of preferential entailment and rational closure.

Definition 45 (Preferential entailment). $(X \rightsquigarrow Y)$ is preferentially entailed by a base $\mathcal{K}$ when every ranked model of $\mathcal{K}$ (i.e. ranked model satisfying all conditionals in $\mathcal{K}$) also satisfies $(X \rightsquigarrow Y)$.

Example 18. In Ex. 17, $\mathcal{K}_\Delta$ also preferential entails $(\mathbf{b} \rightsquigarrow \neg \mathbf{p})$ (“bird is normally not a penguin”). Indeed, if a minimal $\mathbf{b}$ -world satisfied $\mathbf{p}$, then, by the first two conditionals, it could not be a minimal $\mathbf{p}$ -world. Hence there would exist a minimal $\mathbf{p}$ -world below it, which, by the third conditional, would also satisfy $\mathbf{b}$, contradicting the assumption that the original world was a minimal $\mathbf{b}$ -world. At the same time, $\mathcal{K}_\Delta$ does not preferentially entail $(\mathbf{b} \wedge \mathbf{g}) \rightsquigarrow \mathbf{f}$, as it is false in the following model of $\mathcal{K}_\Delta$: $\langle W, \succeq, V \rangle$, where

- $W = \{a, b\}$;
- $w_1 \succeq w_2$ iff $w_1 = a$ and $w_2 = b$;
- $V(a) = \{\mathbf{b}, \mathbf{f}\}$ and $V(b) = \{\mathbf{b}, \mathbf{g}\}$

Giordano et al. [GGOP15] proposed a semantic characterization of rational closure in this setting by introducing a natural ordering on ranked models (comparing ranks world-wise) and capturing rational closure via minimal ranked models, realizing all valuations compatible with the base.

¹In general, there can be no maximal length of a descending chain, even for smooth models. In such a case, we consider the rank of a world to be equal to ∞ . This situation, however, will not occur for specific models of finite bases that we will use in our proof-theoretic study.

Definition 46 (Canonical minimal ranked models). *Let Ξ be a finite set of propositional atoms. A Boolean valuation ρ on Ξ is called compatible with a base $\mathcal{K}$ w.r.t. Ξ when there is no propositional formula A over Ξ such that $\rho \models A$ and for every model M of $\mathcal{K}$, $M \models A \rightsquigarrow \perp$. A ranked model $M = \langle W, \preceq, V \rangle$ is called canonical for $\mathcal{K}$ w.r.t. Ξ when for every valuation ρ compatible with $\mathcal{K}$ w.r.t. Ξ there is $w \in W$ such that $V(w) = \rho$. A ranked model $M = \langle W, \preceq, V \rangle$ is FIMS-minimal for a base $\mathcal{K}$ if $M \models \mathcal{K}$ and there is no ranked model $M' = \langle W, \preceq', V \rangle$ satisfying $\mathcal{K}$ such that $\mathcal{R}^{M'}(w) \leq \mathcal{R}^M(w)$ for every $w \in W$, with strict inequality for some $w \in W$.*

Remark 13. [GGOP15] actually studies two variants of minimal model semantics: fixed interpretations minimal semantics (FIMS), which compares models with the same worlds and valuations while minimizing the ranks, and variable interpretations minimal semantics (VIMS), which also allows valuations to vary. The FIMS variant is more convenient for our proof-theoretic purposes: we will define the rules for reasoning about the relative ranks of worlds with given propositional valuations, rather than about changes to the valuations themselves.

As shown in [GGOP15], canonical FIMS-minimal models exist for all satisfiable finite bases and satisfy exactly the conditionals in the rational closure; moreover, formula ranks coincide with exceptionality ranks, closely connecting the two perspectives.

Theorem 53. ([GGOP15]) *Let Ξ contain all propositional atoms occurring in a base $\mathcal{K}$ and a conditional $(X \rightsquigarrow Y)$. For every canonical and FIMS-minimal ranked model M of $\mathcal{K}$ w.r.t. Ξ ,*

- $\mathcal{R}^M(A) = \mathcal{R}_{\mathcal{K}}^{ex}(A)$ for any Boolean formula A over Ξ ,
- $(X \rightsquigarrow Y) \in RC(\mathcal{K})$ iff $X \rightsquigarrow Y$ is satisfied in M .

7.4 Hypersequent calculus for saturated reasoning in ranked models

We begin the proof-theoretic investigation of rational closure by introducing the basic calculus for reasoning about ranked models of a given conditional knowledge base. We will follow the general approach of [GGOP09] pursued also in Chapters 3 and 4 that characterizes the preference relation in the model via an additional modality $\boxdot$. In [GGOP09] this approach was already applied to KLM logics **P** and **R** (KLM logic **P** plus rational monotonicity), already providing a way to reason about ranked models and preferential entailment. Here however we aim at a calculus that derives the full information about comparative ranks of all antecedents in the knowledge base. This information is essential for subsequent representation of model minimality condition as a rule to characterize rational closure. To achieve this we again employ hypersequents,

Concluding rule:

$$\frac{}{A, \Gamma \vdash A, \Delta \mid \mathcal{H}} (id^K) \quad \frac{}{\boxdot A, \Gamma \vdash_K \boxdot A, \Delta \mid \mathcal{H}} (id_{\boxdot}^{\mathcal{H}})$$

Logical rules:

$$\begin{array}{c} \frac{\neg A, \Gamma \vdash \boxed{A}, \Delta \mid \mathcal{H}}{\neg A, \Gamma \vdash \Delta \mid \mathcal{H}} (\neg L^K) \quad \frac{\boxed{A}, \Gamma \vdash \neg A, \Delta \mid \mathcal{H}}{\Gamma \vdash \neg A, \Delta \mid \mathcal{H}} (\neg R^K) \\[10pt] \frac{\boxed{A, B}, A \wedge B, \Gamma \vdash \Delta \mid \mathcal{H}}{A \wedge B, \Gamma \vdash \Delta \mid \mathcal{H}} (\wedge L^K) \quad \frac{\Gamma \vdash \boxed{A}, A \wedge B, \Delta \mid \mathcal{H} \quad \Gamma \vdash \boxed{B}, A \wedge B, \Delta \mid \mathcal{H}}{\Gamma \vdash A \wedge B, \Delta \mid \mathcal{H}} (\wedge R^K) \\[10pt] \frac{\boxed{A}, A \vee B, \Gamma \vdash \Delta \mid \mathcal{H} \quad \boxed{B}, A \vee B, \Gamma \vdash \Delta \mid \mathcal{H}}{A \vee B, \Gamma \vdash \Delta \mid \mathcal{H}} (\vee L^K) \quad \frac{\Gamma \vdash \boxed{A, B}, A \vee B, \Delta \mid \mathcal{H}}{\Gamma \vdash A \vee B, \Delta \mid \mathcal{H}} (\vee R^K) \\[10pt] \frac{A \rightarrow B, \Gamma \vdash \boxed{A}, \Delta \mid \mathcal{H} \quad \boxed{B}, A \rightarrow B, \Gamma \vdash \Delta \mid \mathcal{H}}{A \rightarrow B, \Gamma \vdash \Delta \mid \mathcal{H}} (\rightarrow L^K) \quad \frac{\boxed{A}, \Gamma \vdash \boxed{B}, A \rightarrow B, \Delta \mid \mathcal{H}}{\Gamma \vdash A \rightarrow B, \Delta \mid \mathcal{H}} (\rightarrow R^K) \end{array}$$

Rank rules:

$$\begin{array}{c} \frac{\boxdot A \vdash_K A \mid \Gamma \vdash_K \boxdot A, \Delta \mid \mathcal{H}}{\Gamma \vdash_K \boxdot A, \Delta \mid \mathcal{H}} (\boxdot_R R^{\mathcal{H}}) \\[10pt] \frac{\boxed{B}, \boxdot A, \Gamma \vdash_K \Delta \mid \boxdot B, \Sigma \vdash_K \boxdot A, \Pi \mid \mathcal{H}}{\boxdot A, \Gamma \vdash_K \Delta \mid \boxdot B, \Sigma \vdash_K \boxdot A, \Pi \mid \mathcal{H}} (\boxdot_R L^{\mathcal{H}}) \\[10pt] \frac{\boxed{\Sigma}, \Gamma \vdash_K \Delta \mid \Sigma \vdash_K \Pi \mid \mathcal{H} \quad \Gamma \vdash_K \Delta \mid \boxed{\Gamma}, \Sigma \vdash_K \Pi \mid \mathcal{H}}{\Gamma \vdash_K \Delta \mid \Sigma \vdash_K \Pi \mid \mathcal{H}} (com_{\mathcal{R}}^{\mathcal{H}}) \end{array}$$

Base rule (applicable for X, Y such that $(X \rightsquigarrow Y) \in \mathcal{K}$):

$$\frac{\boxed{Y}, \boxdot \neg X, X, \Gamma \vdash_K \Delta \mid \mathcal{H}}{\boxdot \neg X, X, \Gamma \vdash_K \Delta \mid \mathcal{H}} (\rightsquigarrow_K^{\mathcal{H}})$$

Analytic cut (applicable for Z such that either $Z \in \bigcup_{(X \rightsquigarrow Y) \in \mathcal{K}} \{X, \boxdot \neg X\}$ or Z appears in the conclusion):

$$\frac{\boxed{Z}, \Gamma \vdash_K \Delta \mid \mathcal{H} \quad \Gamma \vdash_K \boxed{Z}, \Delta \mid \mathcal{H}}{\Gamma \vdash_K \Delta \mid \mathcal{H}} (cut_{an}^{\mathcal{H}})$$

Figure 7.1: Rules of the basic calculus $\mathcal{HS}_{\mathcal{R}}$. $\Gamma^{\mathbb{B}} = \{A \in \Gamma \mid A \text{ is a Boolean formula}\}$, $\Gamma^{\boxdot} = \{\boxdot A \mid \boxdot A \in \Gamma\}$. The addition in each premise w.r.t. the conclusion is highlighted.

together with *analytic cuts*, i.e. restricted versions of the cut rule that are applied only to the formulas present in the conclusion.

To simplify the calculus, we will not consider nesting or Boolean combinations of $\boxdot$ modalities, and the formulas in the derivations will have either the form A or $\boxdot A$, where A is a Boolean formula (we denote the language of such formulas by $\mathcal{L}_1^{\boxdot}$). Evaluation of a formula $\boxdot A$ in a ranked model is standard; it applies to all strictly smaller ranks:

$$M, w \models \boxdot A \quad \text{when } M, u \models A \text{ for all } u \in W \text{ s.t. } \mathcal{R}^M(u) < \mathcal{R}^M(w)$$

Our calculus will employ hypersequents of the form $\Gamma_1 \vdash_{\mathcal{K}} \Delta_1 \mid \dots \mid \Gamma_n \vdash_{\mathcal{K}} \Delta_n$, where Γ_i and Δ_i are multisets of formulas from $\mathcal{L}_1^{\boxdot}$, and $\mathcal{K}$ is a given base. Note that $\mathcal{K}$ is the same in all components $\Gamma_i \vdash_{\mathcal{K}} \Delta_i$. Such hypersequents can be interpreted w.r.t. ranked models using the following disjunctive reading.

Definition 47. $\mathcal{H} = \Gamma_1 \vdash_{\mathcal{K}} \Delta_1 \mid \dots \mid \Gamma_n \vdash_{\mathcal{K}} \Delta_n$ is validated by a ranked model M of the base $\mathcal{K}$ when there exists at least one component $\Gamma_k \vdash_{\mathcal{K}} \Delta_k$ such that every world in M satisfying all formulas in Γ_k also satisfies at least one formula in Δ_k . Conversely, $\mathcal{H}$ is falsified by M when, for each component $\Gamma_k \vdash_{\mathcal{K}} \Delta_k$, there exists a world in M that satisfies all formulas in Γ_k and falsifies all formulas in Δ_k . The hypercontext $\mathcal{H}$ is valid w.r.t. ranked models when it is validated by all ranked models of $\mathcal{K}$.

Our calculus $\mathcal{HS}_{\mathcal{R}}$ for ranked models is given in Fig. 7.1. Concluding and logical rules are standard LK rules in the hypersequent Kleene form (see Sec. 2.3). All other rules add new information to premise(s), either by adding a new component to the hypersequent or by adding formulas to existing components (so the rules are in the Kleene form). Notice that the base $\mathcal{K}$ is the same in the premises and in the conclusion, so it never changes within one derivation. Rule $(\rightsquigarrow_{\mathcal{K}}^{\mathcal{H}})$ applies the satisfiability of a conditional from the base $\mathcal{K}$. Rule $(\boxdot_{\mathcal{R}} L^{\mathcal{H}})$ applies $\boxdot$ -formula to a component corresponding to a world with a strictly smaller rank. Rule $(\boxdot_{\mathcal{R}} R^{\mathcal{H}})$ is a variation of the **GL**-rule for $\boxdot$ -modality, which does not preserve any formulas on the LHS, since it is not necessary in the presence of analytic cuts. Rule $(com_{\mathcal{R}}^{\mathcal{H}})$ reflects the totality of the preference relation in ranked models. Finally, $(cut_{an}^{\mathcal{H}})$ is a version of the standard sequent cut rule, restricted to the formulas in the conclusion, antecedents of the conditionals in the base, and the $\boxdot$ -formulas assessing their ranks. The semantic justification for all these rules comprises the following soundness proof.

Theorem 54 (Soundness of $\mathcal{HS}_{\mathcal{R}}$). *If $\mathcal{H}$ is derivable in $\mathcal{HS}_{\mathcal{R}}$ then $\mathcal{H}$ is valid w.r.t. ranked models.*

Proof. We use the standard contrapositive argument for calculi in Kleene form. For every rule, we show that if the conclusion is falsified by some model M then one of the premises is also falsified by M . This is enough to obtain soundness: if $\mathcal{H}$ was not valid and was derivable at the same time, applying this step bottom-up, we would reach a contradictory

result that the conclusion of a concluding rule is not valid. For logical rules and $(cut_{an}^{\mathcal{H}})$, it is trivial. For $(\rightsquigarrow_K^{\mathcal{H}})$, there is w in M satisfying X and $\Box\neg X$, so w is minimal for X in M and therefore $M, w \models Y$ since $M \models X \rightsquigarrow Y$. For $(com_{\mathcal{R}}^{\mathcal{H}})$, there are w_1 and w_2 in M falsifying the first and the second component of the conclusion, respectively, and either $\mathcal{R}^M(w_1) \leq \mathcal{R}^M(w_2)$, implying that w_1 satisfies formulas in $\Sigma^{\Box}$, or $\mathcal{R}^M(w_2) \leq \mathcal{R}^M(w_1)$, implying that w_2 satisfies formulas in $\Gamma^{\Box}$. For $(\Box_{\mathcal{R}}R^{\mathcal{H}})$, there is w in M such that $M, w \not\models \Box A$, so there should be w' in M with $\mathcal{R}^M(w') = \mathcal{R}^M(\neg A) < \mathcal{R}^M(w)$ and $M, w' \models \Box A$ and $M, w' \not\models A$, which falsifies the additional component of the premise. For $(\Box_{\mathcal{R}}L^{\mathcal{H}})$, there are w_1 and w_2 in M falsifying the first and the second component of the conclusion respectively, so $\mathcal{R}^M(w_1) \leq \mathcal{R}^M(\neg A) < \mathcal{R}^M(w_2)$ and $M, w_2 \models \Box B$, so $M, w_1 \models B$. $\square$

Despite the presence of restricted form of cut rule, the calculus $\mathcal{HS}_{\mathcal{R}}$ is still analytic, since the formulas added to the premises come from formulas already present in the conclusion. More specifically, $\mathcal{HS}_{\mathcal{R}}$ enjoys the relaxed subformula property in the following form.

Proposition 2. *In any derivation in $\mathcal{HS}_{\mathcal{R}}$ with $\mathcal{H} = \Gamma_1 \vdash_{\mathcal{K}} \Delta_1 \mid \dots \mid \Gamma_n \vdash_{\mathcal{K}} \Delta_n$ in the root, each hypersequent contains only subformulas of formulas from set*

$$\mathcal{F}(\mathcal{H}) = \bigcup_{1 \leq i \leq n} (\Gamma_i \cup \Delta_i) \cup \bigcup_{(A \rightsquigarrow B) \in \mathcal{K}} \{\Box\neg A, B\}.$$

This implies that the amount of information about the model that can be added to hypersequent within one derivation is limited. Therefore the size of derivations can also be limited, unless they contain redundant rule applications.

Definition 48. *We say that a hypersequent component $\Gamma \vdash_{\mathcal{K}} \Delta$ is subsumed by another component $\Gamma' \vdash_{\mathcal{K}} \Delta'$ if all formulas appearing in Γ also appear in Γ' and all formulas appearing in Δ also appear in Δ' , and that hypersequent $\mathcal{H}$ is subsumed by another hypersequent $\mathcal{H}'$ if every component of $\mathcal{H}$ is subsumed by some component of $\mathcal{H}'$. We call a rule application redundant when one of the premises is subsumed by the conclusion. A derivation without redundant rule applications is called non-redundant.*

Redundant rule applications are not necessary, since any derivation of the subsumed premise can be performed on the conclusion already (since all rules are stable under weakening). Since the number of formulas we can add to one component is limited by $|\mathcal{F}(\mathcal{H})|$ and the number of different components is also limited, every branch in a non-redundant derivation arrives after a limited number of steps either to an application of a concluding rule or to a hypersequent for which only redundant rule applications are possible (we call such hypersequent *saturated*).

Lemma 45. *In any non-redundant derivation of a hypersequent $\mathcal{H}$ in $\mathcal{HS}_{\mathcal{R}}$, every branch contains $\mathcal{O}((|\mathcal{F}(\mathcal{H})| + |\mathcal{H}|)^2)$ hypersequents, where $|\mathcal{H}|$ is number of components in $\mathcal{H}$.*

Proof. Rule $(\boxdot_{\mathcal{R}} R^{\mathcal{H}})$, which introduces a new component, can be applied non-redundantly at most once in every branch for every $\boxdot A \in \mathcal{F}(\mathcal{H})$, so hypersequents in every branch have $|\mathcal{H}| + |\mathcal{F}(\mathcal{H})|$ components and a non-redundant application of a rule reduces the number of subformulas from $\mathcal{F}(\mathcal{H})$ not occurring in some component. $\square$

The depth of non-redundant derivation in $\mathcal{HS}_{\mathcal{R}}$ is therefore polynomially bounded. Note that the depth of **LK**-subderivations is also polynomially bounded. Accordingly, the naive proof search in $\mathcal{HS}_{\mathcal{R}}$ can be performed by performing non-redundant rule applications in an arbitrary order until every branch is either closed or reaches saturation. This proof search is exhaustive and does not require backtracking in rule applications, since rules in Kleene form are always *invertible*, meaning that if any premise is not derivable, then the conclusion is not derivable either (which is true again due to weakening). This procedure allows to non-deterministically verify non-derivability in polynomial time (matching *co-NP*-completeness of preferential entailment checking), by guessing the branch leading to a non-derivable saturated hypersequent and verifying that no further non-redundant rule applications are possible.

Finally, we can reconstruct a ranked model falsifying $\mathcal{H}$ from any non-derivable saturated hypersequent subsuming $\mathcal{H}$, which simultaneously gives us a procedure for countermodel extraction from failed derivation in $\mathcal{HS}_{\mathcal{R}}$ and provides a semantic proof of completeness of $\mathcal{HS}_{\mathcal{R}}$.

Theorem 55 (Completeness of $\mathcal{HS}_{\mathcal{R}}$). *If hypersequent $\mathcal{H}$ is valid w.r.t. ranked models then it is derivable in $\mathcal{HS}_{\mathcal{R}}$.*

Proof. By contraposition: assuming $\mathcal{H}$ is not derivable, we first show that there exists underivable saturated hypersequent subsuming $\mathcal{H}$ and then use it to build a model of $\mathcal{K}$ falsifying $\mathcal{H}$. Saturated hypersequent can be obtained by applying naive proof search: performing arbitrary non-redundant rule applications to $\mathcal{H}$ while it is possible results in a derivation with at least one leaf labeled by a saturated and non-derivable hypersequent $\mathcal{H}'$ (otherwise derivation can be closed), which subsumes $\mathcal{H}$. Now, we consider the model $M = \langle W, \preceq, V \rangle$ where

- W is the set of components in $\mathcal{H}'$, not subsumed by any other component in $\mathcal{H}'$,
- $V(\Gamma \vdash_{\mathcal{K}} \Delta) = \{p \in \text{Atm} \mid p \in \Gamma\}$,
- $(\Gamma \vdash_{\mathcal{K}} \Delta) \preceq (\Sigma \vdash_{\mathcal{K}} \Pi)$ iff $\Gamma^{\boxdot} \subseteq \Delta^{\boxdot}$.

Note that preference relation $\preceq$ is reflexive, transitive and total (due to saturation w.r.t. $(\text{com}_{\mathcal{R}}^{\mathcal{H}})$), and also trivially satisfies smoothness since W is finite, so M is a ranked model. We now show that every component $\Gamma \vdash_{\mathcal{K}} \Delta$ in W satisfies in M all formulas in Γ and falsifies all formulas in Δ . For Boolean formulas, this follows immediately from the saturation w.r.t. logical rules and non-derivability by concluding rules. If $\boxdot B \in \Gamma$

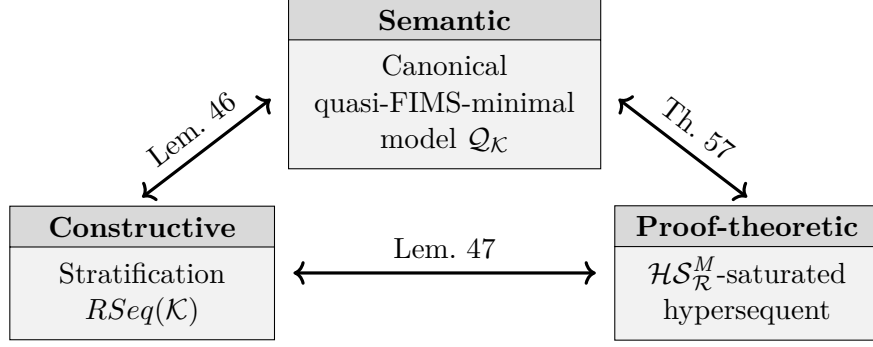


Figure 7.2: Correspondence between non-entailment witnesses w.r.t. different characterizations of rational closure.

7.5 Hypersequent calculus for rational closure

We now extend our calculus to enable reasoning about ranks in minimal ranked models. To preserve the exact connection with the semantics, we, however need to slightly relax the notion of FIMS-minimal model to allow worlds with the same valuation on different layers, while preserving the rank for each valuation. This extra degree of freedom is reflected in the following definition of *quasi-FIMS-minimal model*.

Definition 49. *Ranked model M is called quasi-FIMS-minimal for $\mathcal{K}$ w.r.t. Ξ if there is a model M' FIMS-minimal for $\mathcal{K}$ w.r.t. Ξ , such that $\mathcal{R}^M(\rho) = \mathcal{R}^{M'}(\rho)$ for each $\rho \in 2^{Atm}$.*

Since all FIMS-minimal and quasi-FIMS-minimal models have the same ranks of valuations, they are equivalent w.r.t. ranks of formulas and satisfied conditionals. However, this subtle formal distinction is important for our proof-theoretic characterisation, since copies of the same world on different layers may satisfy different $\Box$ -formulas. E.g. $\Box\perp$ is valid in all FIMS-minimal models of the empty base, but not in all quasi-FIMS-minimal models of it.

The calculus $\mathcal{HS}_{\mathcal{R}}^M$ for reasoning in canonical (quasi-)FIMS-minimal models extends $\mathcal{HS}_{\mathcal{R}}$ with the rules of the antisequent calculus *antiLK* (Fig. 2.6,p.26) and the following rule with an anti-sequent premise:

$$\frac{\overline{(\mathcal{K} \setminus \Delta^{\rightsquigarrow})} \not\models B}{\Box A, \Gamma \vdash_{\mathcal{K}} \Delta \mid \Box B, \Sigma \vdash_{\mathcal{K}} \Box A, \Pi \mid \mathcal{H}} (M)$$

where $\Delta^{\rightsquigarrow} = \{(X \rightsquigarrow Y) \in \mathcal{K} \mid \Box \neg X \in \Delta\}$, and $\overline{(\mathcal{K} \setminus \Delta^{\rightsquigarrow})}$ is the materialization of $(\mathcal{K} \setminus \Delta^{\rightsquigarrow})$, i.e. $\overline{(\mathcal{K} \setminus \Delta^{\rightsquigarrow})} = \{A \rightarrow B \mid (A \rightsquigarrow B) \in \mathcal{K} \setminus \Delta^{\rightsquigarrow}\}$.

Note that the conclusion of rule (M) is the same as in the rule $(\Box_{\mathcal{R}}L^{\mathcal{H}})$, so it encodes some statement about worlds w_1 and w_2 with $\mathcal{R}^M(w_1) < \mathcal{R}^M(w_2)$. Intuitively, the rule encodes the fact that in a canonical quasi-FIMS-minimal model for $\mathcal{K}$, formula $\neg B$ should

be classically inconsistent with the constraints that $\mathcal{K}$ imposes on any ranks below the rank of $\neg B$. More formally, it is demonstrated in the following soundness proof.

Theorem 56 (Soundness of $\mathcal{HS}_{\mathcal{R}}^M$). *If $\mathcal{H}$ is derivable in $\mathcal{HS}_{\mathcal{R}}^M$ then it is valid w.r.t. canonical quasi-FIMS-minimal ranked models.*

Proof. We follow the same inductive proof as in Th. 54, soundness of all rules but (M) is shown there. For (M), we assume that the conclusion is falsified in some canonical quasi-FIMS-minimal model M of $\mathcal{K}$ and that simultaneously antisequent in the premise is valid, i.e. that $\neg B$ is classically consistent with $(\mathcal{K} \setminus \Delta_{\mathcal{K}}^{\sim})$, and we will show that it leads to a contradiction. Due to falsification of the conclusion, there are two worlds w_1 and w_2 in M corresponding to the first and the second component of the conclusion. Let $r = \mathcal{R}^M(w_1)$. For all $\Box \neg X \in \Delta$, $\mathcal{R}_{\mathcal{K}}^{ex}(X) = \mathcal{R}^M(X) < r$, so $\mathcal{C}_r \subseteq \mathcal{K} \setminus \Delta_{\mathcal{K}}^{\sim}$. The validity of the antisequent in the premise therefore implies also that $\neg B$ is classically consistent with $\mathcal{C}_r$, so $\mathcal{R}^M(\neg B) = \mathcal{R}_{\mathcal{K}}^{ex}(\neg B) \leq r$. At the same time, $r = \mathcal{R}^M(w_1) \leq \mathcal{R}^M(\neg A) < \mathcal{R}^M(w_2) \leq \mathcal{R}^M(\neg B)$, leading to a contradiction. $\square$

Note the crucial feature of the rule (M): the validity of its antisequent premise can be violated by extension of $\mathcal{K}$, but is preserved under addition of arbitrary formulas to components, since more formulas in Δ leads to fewer formulas in the antisequent. Therefore derivability in $\mathcal{HS}_{\mathcal{R}}^M$ is not conservative w.r.t. base extension (as needed for global non-monotonicity), while simultaneously conservative w.r.t. weakening of the hypersequent (as needed for invertibility, which we used for proof search and completeness proof). Since the additional rule (M) does not have hypersequents as premises it does not affect our arguments about analyticity and redundant rule applications. So Prop. 2 and Lem. 45 remain true for $\mathcal{HS}_{\mathcal{R}}^M$. The naive proof search procedure can also be performed the same way, applying the rules of $\mathcal{HS}_{\mathcal{R}}$ until branch is closed or saturation is reached, and then attempting all possible applications of (M) (either by performing proof search in *antiLK* or by checking classical entailment in any other way). This, however, leads to higher complexity of verifying non-derivability with this strategy: after guessing the failed branch non-deterministically, checking that the final sequent can not be derived by (M) requires co-non-deterministic polynomial time, leading to a solution in class Σ_2^P . This complexity bound is higher than the optimal P^{NP} complexity of the rational closure, although at the same level of the polynomial hierarchy. At the end of this section, we will see that proof search in $\mathcal{HS}_{\mathcal{R}}^M$ can in fact be performed in P^{NP} using an oracle that computes exceptionality ranks.

We can still use the same strategy for completeness proof, constructing a countermodel based on a saturated hypersequent in a failed branch. However, due to quasi-FIMS-minimality restriction we need a more involved countermodel construction and we need to thread the proof through the constructive characterization of rational closure, establishing three-sided natural connection (shown in Fig. 7.2) between saturated hypersequents w.r.t. $\mathcal{HS}_{\mathcal{R}}^M$, constructive characterization $RSeq(\mathcal{K})$ and quasi-FIMS-minimal model semantics. Specifically, we prove completeness w.r.t. a specific canonical quasi-FIMS-minimal model,

defined on the basis of $RSeq(\mathcal{K})$: at each layer k it contains worlds with all valuations consistent with materialization of the layer $\mathcal{C}_k$ from constructive characterization of rational closure (Def. 43, p.158). The resulting model is equivalent to a canonical model in [BP98] obtained by merging all ranked models of $\mathcal{K}$.

Definition 50. For finite set Ξ of propositional atoms, the target canonical quasi-FIMS-minimal model $\mathcal{Q}_\mathcal{K} = \langle W_\mathcal{K}^\mathcal{Q}, \preceq_\mathcal{K}^\mathcal{Q}, V_\mathcal{K}^\mathcal{Q} \rangle$ of finite $\mathcal{K}$ w.r.t. Ξ is defined as follows:

- $W_\mathcal{K}^\mathcal{Q} = \{ \langle i, \rho \rangle \mid i \in \mathbb{N}, \rho \text{ is a valuation on } \Xi \text{ s.t. } \rho \models A \rightarrow B \text{ for all } (A \rightsquigarrow B) \in \mathcal{C}_i \},$
- $V_\mathcal{K}^\mathcal{Q}(\langle i, \rho \rangle) = \rho,$
- $\langle i, \rho \rangle \preceq_\mathcal{K}^\mathcal{Q} \langle j, \rho' \rangle$ iff $i \leq j$.

Lemma 46. $\mathcal{Q}_\mathcal{K}$ is a canonical quasi-FIMS-minimal model of $\mathcal{K}$ w.r.t. Ξ .

Proof. Preference relation $\preceq_\mathcal{K}^\mathcal{Q}$ is reflexive, transitive, total, and smooth. For every $(A \rightsquigarrow B) \in \mathcal{K}$, $\mathcal{R}^{\mathcal{Q}_\mathcal{K}}(A) = \mathcal{R}_\mathcal{K}^{ex}(A) < \mathcal{R}_\mathcal{K}^{ex}(A \wedge \neg B) = \mathcal{R}^{\mathcal{Q}_\mathcal{K}}(A \wedge \neg B)$ by Def. 42 and Def. 43. So $\mathcal{Q}_\mathcal{K}$ is a ranked model of $\mathcal{K}$. Let us show that it is canonical for $\mathcal{K}$ w.r.t. Ξ . Let ρ be a valuation on Ξ compatible with $\mathcal{K}$ w.r.t. Ξ , let $Z = (\bigwedge_{x \in \rho} x) \wedge (\bigwedge_{x \in \Xi \setminus \rho} \neg x)$ and let M be some canonical quasi-FIMS-minimal model of $\mathcal{K}$ w.r.t. Ξ . Due to canonicity of M , there is a world w_1 in M with valuation ρ , so $M, w_1 \models Z$, so $\mathcal{R}_\mathcal{K}^{ex}(Z) = \mathcal{R}^M(Z) \neq \infty$, therefore Z is classically consistent with $\overline{\mathcal{C}}_k$ for some $k \in \mathbb{N}$, implying $\langle k, \rho \rangle \in W_\mathcal{K}^\mathcal{Q}$. It remains to show that it is quasi-FIMS-minimal for $\mathcal{K}$ w.r.t. Ξ . If not, then there is a valuation ρ on Ξ such that $\mathcal{R}^{\mathcal{Q}_\mathcal{K}}(\rho) > \mathcal{R}^M(\rho)$ for some FIMS-minimal model M of $\mathcal{K}$. Then there is $(A \rightsquigarrow B) \in \mathcal{C}_{\mathcal{R}^M(\rho)}$ such that $\rho \not\models A \rightarrow B$. Then $\mathcal{R}^M(A) < \mathcal{R}^M(A \wedge \neg B) \leq \mathcal{R}^M(\rho)$. But also $\mathcal{R}_\mathcal{K}^{ex}(A) \geq \mathcal{R}^M(\rho)$ since $(A \rightsquigarrow B) \in \mathcal{C}_{\mathcal{R}^M(\rho)}$ so A is exceptional for $\mathcal{C}_i$ for $i < \mathcal{R}^M(\rho)$. We thus get a contradiction with Th. 53 (p.160). $\square$

In turn, saturated non-derivable hypersequents in $\mathcal{HS}_\mathcal{R}^M$ are closely connected to the constructive characterization $RSeq(\mathcal{K})$: every such hypersequent $\mathcal{H}$ contains a certain representation of a prefix of $RSeq(\mathcal{K})$. To see this, we define a ranking for non-subsumed (within $\mathcal{H}$) components of $\mathcal{H}$. Consider the following relation on non-subsumed components of $\mathcal{H}$:

$$(\Gamma \vdash_\mathcal{K} \Delta) \sqsubset^\uparrow (\Sigma \vdash_\mathcal{K} \Pi) \quad \text{when } \Gamma^\boxminus \subseteq \Sigma^\boxminus \text{ and } \Delta^\boxminus \cap \Sigma^\boxminus \neq \emptyset.$$

Let rank $R^\mathcal{H}(\Gamma, \Delta)$ be the maximal length of chain $(\Gamma \vdash_\mathcal{K} \Delta) \sqsubset^\uparrow \dots \sqsubset^\uparrow (\Gamma_n \vdash_\mathcal{K} \Delta_n)$ in $\mathcal{H}$. $R^\mathcal{H}(\Gamma, \Delta)$ is well-defined since $\mathcal{H}$ is finite, while $\sqsubset^\uparrow$ is transitive and irreflexive due to saturation w.r.t. $(id_\boxminus^\mathcal{H})$. Now consider the subset of $\mathcal{K}$ corresponding to a non-subsumed component $\Gamma \vdash_\mathcal{K} \Delta$:

$$\Gamma^\rightsquigarrow = \{ (A \rightsquigarrow B) \in \mathcal{K} \mid \boxminus \neg A \in \Gamma \}$$

(it is the same notion we use in (M), but for LHS of the component). We can show that each such set corresponds to an element of $RSeq(\mathcal{K})$ with index $R^\mathcal{H}(\Gamma, \Delta)$.

Lemma 47. *If $\Gamma \vdash_{\mathcal{K}} \Delta$ is a non-subsumed component in a hypersequent $\mathcal{H}$, saturated and non-derivable in $\mathcal{HS}_{\mathcal{R}}^M$, then $\Gamma^{\rightsquigarrow} = \mathcal{C}_{R^{\mathcal{H}}(\Gamma, \Delta)}$.*

Proof. By induction on $R^{\mathcal{H}}(\Gamma, \Delta)$, which we denote r for brevity.

Suppose $r = 0$. Then for every $(A \rightsquigarrow B) \in \mathcal{K}$, $\Box \neg A \in \Gamma$, since otherwise $\Box \neg A \in \Delta$ due to saturation w.r.t. $(cut_{an}^{\mathcal{H}})$, and due to saturation w.r.t. $(\Box_{\mathcal{R}} R^{\mathcal{H}})$ there would be a component $\Sigma \vdash_{\mathcal{K}} \Delta$ such that $\Box \neg A \in \Sigma$, so $\Gamma^{\Box} \subseteq \Sigma$ due to saturation w.r.t. $(com_{\mathcal{R}}^{\mathcal{H}})$, implying $(\Gamma \vdash_{\mathcal{K}} \Delta) \sqsubset^{\uparrow} (\Sigma \vdash_{\mathcal{K}} \Delta)$, which contradicts $r = 0$. Therefore $\Gamma^{\rightsquigarrow} = \mathcal{K} = \mathcal{C}_0$.

Now suppose $r > 0$. Then there is a non-subsumed component $\Sigma \vdash_{\mathcal{K}} \Pi$ in $\mathcal{H}$, such that $(\Gamma \vdash_{\mathcal{K}} \Delta) \sqsubset^{\uparrow} (\Sigma \vdash_{\mathcal{K}} \Pi)$ and $R^{\mathcal{H}}(\Sigma, \Pi) = r - 1$. By inductive hypothesis $\Sigma^{\rightsquigarrow} = \mathcal{C}_{r-1}$. For any $(A \rightsquigarrow B) \in \Gamma^{\rightsquigarrow}$, A is classically inconsistent with $\overline{\Sigma^{\rightsquigarrow}} = \overline{\mathcal{C}_{r-1}}$ due to non-derivability via (M) and saturation w.r.t. $(cut_{an}^{\mathcal{H}})$. So it remains to show that for any $(A \rightsquigarrow B) \in \Sigma^{\rightsquigarrow} \setminus \Gamma^{\rightsquigarrow}$, A is classically consistent with $\overline{\mathcal{C}_{r-1}}$. In such case, $\Box \neg A \in \Delta$ due to saturation w.r.t. $(cut_{an}^{\mathcal{H}})$, so due to saturation w.r.t. $(\Box_{\mathcal{R}} R^{\mathcal{H}})$ there is a non-subsumed component $\Theta \vdash_{\mathcal{K}} \Xi$ such that $\{\Box \neg A, A\} \subseteq \Theta$ and $\Gamma^{\Box} \subseteq \Theta$ (due to saturation w.r.t. $(com_{\mathcal{R}}^{\mathcal{H}})$), so $(\Gamma \vdash_{\mathcal{K}} \Pi) \sqsubset^{\uparrow} (\Theta \vdash_{\mathcal{K}} \Xi)$, which means that $R^{\mathcal{H}}(\Theta, \Xi) < r$. A is classically consistent with $\overline{\Theta^{\rightsquigarrow}}$: for every $(X \rightsquigarrow Y) \in \Theta^{\rightsquigarrow}$ either $X \in \Xi$ or $Y \in \Theta$ due to saturation w.r.t. $(cut_{an}^{\mathcal{H}})$ and $(\rightsquigarrow_{\mathcal{K}}^{\mathcal{H}})$, and, as usual for saturated hypersequents, there is a valuation satisfying all Boolean formulas in Θ and falsifying all Boolean formulas in Ξ due to non-derivability and saturation w.r.t. logical rules of $\mathcal{HS}_{\mathcal{R}}^M$ (specifically, valuation $\{p \in Atm \mid p \in \Theta\}$). Therefore A is classically consistent with $\overline{\mathcal{C}_k}$ for some $k < r$, which means it is also consistent with $\overline{\mathcal{C}_{r-1}}$. $\square$

Using Lem. 47, we can now prove that each underivable saturated hypersequent for base $\mathcal{K}$ is falsified by the target canonical quasi-FIMS-minimal model $\mathcal{Q}_{\mathcal{K}}$, establishing completeness of $\mathcal{HS}_{\mathcal{R}}^M$ w.r.t. canonical quasi-FIMS-minimal model semantics.

Theorem 57 (Completeness of $\mathcal{HS}_{\mathcal{R}}^M$). *If $\mathcal{H}$ is valid w.r.t. canonical quasi-FIMS-minimal ranked models then it is derivable in $\mathcal{HS}_{\mathcal{R}}^M$.*

Proof. Analogously to the proof of Th. 55, if $\mathcal{H}$ is non-derivable in $\mathcal{HS}_{\mathcal{R}}^M$ we can obtain by naive proof search a saturated $\mathcal{H}'$ non-derivable in $\mathcal{HS}_{\mathcal{R}}^M$ and subsuming $\mathcal{H}$. Now we will show that every non-subsumed component $\Gamma \vdash_{\mathcal{K}} \Delta$ in $\mathcal{H}'$ is falsified in $\mathcal{Q}_{\mathcal{K}}$ in the world $w^* = \langle k^*, \rho^* \rangle$ for some valuation ρ^* . Take $k^* = R^{\mathcal{H}'}(\Gamma, \Delta)$ and take $\rho^* = \{p \in Atm \mid p \in \Gamma\}$. ρ^* satisfies all Boolean formulas in Γ and falsifies all Boolean formulas in Δ due to non-derivability and saturation w.r.t. logical rules. For every $(A \rightsquigarrow B) \in \Gamma^{\rightsquigarrow}$, either $A \in \Delta$ or $B \in \Gamma$ due to saturation w.r.t. $(cut_{an}^{\mathcal{H}})$ and $(\rightsquigarrow_{\mathcal{K}}^{\mathcal{H}})$, so $\rho^* \models (A \rightarrow B)$, so due to Lem. 47, $w^* \in W_{\mathcal{K}}^{\mathcal{Q}}$.

Let us prove by induction on $R^{\mathcal{H}'}(\Gamma, \Delta)$ that w^* indeed falsifies $\Gamma \vdash_{\mathcal{K}} \Delta$. If $R^{\mathcal{H}'}(\Gamma, \Delta) = 0$ then w^* trivially satisfies the formulas in $\Gamma^{\Box}$ and $\Delta^{\Box} = \emptyset$, as demonstrated in the proof of Lem. 47. Now suppose $R^{\mathcal{H}'}(\Gamma, \Delta) > 0$. Then there is a component $\Sigma \vdash_{\mathcal{K}} \Pi$

such that $(\Gamma \vdash_{\mathcal{K}} \Delta) \sqsubset^{\uparrow} (\Sigma \vdash_{\mathcal{K}} \Pi)$ and $\Sigma^{\rightsquigarrow} = \mathcal{C}_{k^*-1}$. For every $\boxplus A \in \Gamma$, $\neg A$ is classically inconsistent with $\Sigma^{\rightsquigarrow}$ due to underderivability via (M) and saturation w.r.t. $(cut_{an}^{\mathcal{H}})$, so for any world $\langle k, \rho \rangle \in \mathcal{Q}_{\mathcal{K}}$ if $k \leq k^* - 1$ then ρ does not satisfy $\neg A$, implying $\mathcal{R}^{\mathcal{Q}_{\mathcal{K}}}(\neg A) \geq k^* = \mathcal{R}^{\mathcal{Q}_{\mathcal{K}}}(w^*)$, implying $\mathcal{Q}_{\mathcal{K}}, w^* \models \boxplus A$. For every $\boxplus A \in \Delta$, due saturation w.r.t. $(\boxplus_{\mathcal{R}} R^{\mathcal{H}})$ there is a component $\Theta \vdash_{\mathcal{K}} \Xi$ such that $\boxplus A \in \Theta$ and $A \in \Xi$, implying that $(\Gamma \vdash_{\mathcal{K}} \Delta) \sqsubset^{\uparrow} (\Theta \vdash_{\mathcal{K}} \Xi)$ due to underderivability and saturation w.r.t. $(com_{\mathcal{R}}^{\mathcal{H}})$, so by the inductive hypothesis A is satisfied in some world $w \in \mathcal{Q}_{\mathcal{K}}$ with $\mathcal{R}^{\mathcal{Q}_{\mathcal{K}}}(w) < \mathcal{R}^{\mathcal{Q}_{\mathcal{K}}}(w^*)$, implying $\mathcal{Q}_{\mathcal{K}}, w^* \not\models \boxplus A$ $\square$

We therefore can use $\mathcal{HS}_{\mathcal{R}}^M$ as an analytic proof calculus for rational closure.

Corollary 5. $(C \rightsquigarrow D) \in RC(\mathcal{K})$ iff $\boxplus \neg C, C \vdash_{\mathcal{K}} D$ is derivable in $\mathcal{HS}_{\mathcal{R}}^M$.

Example 20. The conditional *bird* $\wedge$ *green* $\rightsquigarrow$ *flies* from Ex. 17 that belongs to rational closure of $\mathcal{K}_{\Delta}$, but is not preferentially entailed by it, can be inferred in $\mathcal{HS}_{\mathcal{R}}^M$ as follows:

$$\begin{array}{c}
 \frac{\frac{f, \boxplus \neg b, b, g, \boxplus \neg(b \wedge g), b \wedge g \vdash_{\mathcal{K}_{\Delta}} f}{\boxplus \neg b, b, g, \boxplus \neg(b \wedge g), b \wedge g \vdash_{\mathcal{K}_{\Delta}} f} (id^{\mathcal{K}})}{\frac{f, \boxplus \neg b, b, g, \boxplus \neg(b \wedge g), b \wedge g \vdash_{\mathcal{K}_{\Delta}} f}{\boxplus \neg b, b, g, \boxplus \neg(b \wedge g), b \wedge g \vdash_{\mathcal{K}_{\Delta}} f} (\rightsquigarrow_{\mathcal{K}}^{\mathcal{H}})} \quad \frac{\frac{b \rightarrow f, p \rightarrow \neg f, p \rightarrow b \not\models \neg(b \wedge g)}{\boxplus \neg b \vdash_{\mathcal{K}_{\Delta}} \neg b \mid b, g, \boxplus \neg(b \wedge g), b \wedge g \vdash_{\mathcal{K}_{\Delta}} \boxplus \neg b, f} (M)}{\frac{b, g, \boxplus \neg(b \wedge g), b \wedge g \vdash_{\mathcal{K}_{\Delta}} \boxplus \neg b, f}{\boxplus \neg b, b, g, \boxplus \neg(b \wedge g), b \wedge g \vdash_{\mathcal{K}_{\Delta}} f} (\boxplus_{\mathcal{R}} R^{\mathcal{H}})} \\
 \frac{\frac{b, g, \boxplus \neg(b \wedge g), b \wedge g \vdash_{\mathcal{K}_{\Delta}} f}{\boxplus \neg(b \wedge g), b \wedge g \vdash_{\mathcal{K}_{\Delta}} f} (\wedge L^{\mathcal{K}})}{\frac{b, g, \boxplus \neg(b \wedge g), b \wedge g \vdash_{\mathcal{K}_{\Delta}} f}{\boxplus \neg(b \wedge g), b \wedge g \vdash_{\mathcal{K}_{\Delta}} f} (cut_{an}^{\mathcal{H}})}
 \end{array}$$

The antisequent in the leaf is valid (e.g. valuation $\{b, f\}$ satisfies the implications on the LHS, but does not satisfy $\neg(b \wedge g)$), and is therefore derivable in antiLK.

The expressive power of $\mathcal{HS}_{\mathcal{R}}^M$ goes beyond reasoning in rational closure. The calculus can be used to derive arbitrary statements about the ranks of formulas in the canonical quasi-FIMS-minimal models associated with a knowledge base, as long as these statements are expressible via hypersequents with formulas from $\mathcal{L}_1^{\boxplus}$. This additional expressivity does not increase computational complexity: verification of such statements remains in P^{NP} , as demonstrated in the following lemma.

Lemma 48. It can be decided in P^{NP} -time whether $\mathcal{Q}_{\mathcal{K}}$ falsifies a given hypersequent $\mathcal{H}$.

Proof. Let us verify falsification of each component $\Gamma_i \vdash_{\mathcal{K}} \Delta_i$ of $\mathcal{H}$ separately. Let $r = \min\{\mathcal{R}_{\mathcal{K}}^{ex}(\neg A) \mid \boxplus A \in \Gamma_i\}$. Let us show that $\mathcal{Q}_{\mathcal{K}}$ falsifies $\Gamma_i \vdash_{\mathcal{K}} \Delta_i$ iff the following two conditions are simultaneously true:

1. the set of formulas

$$\overline{\mathcal{C}_r} \cup \{B \mid B \text{ is a Boolean formula in } \Gamma_i\} \cup \{\neg B \mid B \text{ is a Boolean formula in } \Gamma_i\}$$

is classically consistent;

2. $\mathcal{R}_{\mathcal{K}}^{ex}(\neg A) < r$ for every $\Box A \in \Delta_i$.

Let $\Gamma_i \vdash_{\mathcal{K}} \Delta_i$ be falsified in some world w in $\mathcal{Q}_{\mathcal{K}}$. $\mathcal{R}^{\mathcal{Q}_{\mathcal{K}}}(w) \leq r$, since every $\Box A \in \Gamma_i$ is satisfied in w . Then valuation of w satisfies all implications in $\overline{\mathcal{C}_r}$ by definition of $\mathcal{Q}_{\mathcal{K}}$, satisfies all Boolean formulas in Γ_i and falsifies all Boolean formulas in Δ_i , thus ensuring the first condition. w also falsifies every $\Box \neg A \in \Delta_i$, implying $\mathcal{R}^{\mathcal{Q}_{\mathcal{K}}}(\neg A) < \mathcal{R}^{\mathcal{Q}_{\mathcal{K}}}(w) \leq r$, ensuring the second condition. Conversely, suppose the two conditions are satisfied, and take an arbitrary valuation ρ satisfying all formulas in the first condition. Then $\langle r, \rho \rangle$ is a world in $\mathcal{Q}_{\mathcal{K}}$, and it falsifies $\Gamma_i \vdash \Delta_i$. For Boolean formulas this is given by the first condition, for $\Box A \in \Delta_i$ - by the second condition, and for $\Box A \in \Gamma_i$ - by the choice of r . Both conditions can be verified in P^{NP} -time by the constructing the stratification of $\mathcal{K}$ and computing exceptionality ranks directly by Def. 42 (p.157). $\square$

Note that since our completeness proof uses the specific target canonical quasi-FIMS-minimal model $\mathcal{Q}_{\mathcal{K}}$ while soundness is established w.r.t. an arbitrary canonical quasi-model, it follows that a hypersequent is valid in all canonical quasi-FIMS-minimal models if and only if it is valid in the target one. Using the validity test from Lem. 48 as an oracle for checking whether a component is valid in $\mathcal{Q}_{\mathcal{K}}$, proof search in $\mathcal{HS}_{\mathcal{R}}^M$ can be performed, in a certain sense, within optimal P^{NP} -time. More precisely, not only can derivability be decided in P^{NP} , but whenever a hypersequent is not derivable, a failing branch ending in a saturated hypersequent can also be constructed within the same complexity bound. This can be achieved by repeatedly performing arbitrary non-redundant rule applications and, at each step, using the oracle to identify a premise falsified in $\mathcal{Q}_{\mathcal{K}}$ and continuing the derivation with that premise. The resulting saturated hypersequent can serve as a syntactic witness of non-validity. However, this proof-search procedure still relies heavily on the computation of exceptionality ranks. It is therefore an interesting open problem whether there exists a more direct complexity-optimal proof-search strategy, one that arises naturally from the calculus itself and does not explicitly rely on the constructive characterization of rational closure.

Conclusion

In this thesis, we developed uniform and modular analytic characterizations for five frameworks of conditional formalisms: Burgess’s Preferential Conditional Logic, Åqvist’s logics, Input/Output logics, the KLM hierarchy of non-monotonic logics, and rational closure. Across these different settings, we introduced modular sequent and hypersequent calculi, new semantic constructions and frame conditions, and systematic methods for reconstructing countermodels from failed derivations. These characterizations were then used to study the considered logics from several complementary perspectives: to establish exact complexity bounds, to obtain embeddings into modal and classical logics, to uncover new structural connections between different families of conditionals, and to support practical automated deduction tools based on SAT and SMT solving. This demonstrates that analytic methods provide a fruitful common perspective on conditional reasoning, allowing techniques developed in one setting to be adapted and transferred to others.

The obtained results naturally suggest a number of further directions. On the one hand, many important conditional formalisms remain outside the scope of the present work, offering exciting opportunities to further extend and adapt the analytic methodology developed here. The flexibility of the approach suggests that it may be successfully modified to accommodate a broader range of logics and conditional systems. On the other hand, even for the systems studied here, a number of proof-theoretic, semantic, computational, and practical questions remain open. In this concluding chapter, we first discuss several related conditional frameworks and consider how the methods introduced here may be applied to them (Sec. 8.1), before outlining additional open problems and directions for future work (Sec. 8.2).

8.1 Related conditional formalisms

The landscape of conditional logics is remarkably rich and diverse, and this thesis inevitably covers only some known formalisms. In this section, we discuss a selection of

related conditional formalisms and the potential applicability of our analytic methodology to them.

8.1.1 Conditional Beliefs

An important interpretation of conditionals, not discussed in this thesis so far, is the doxastic one, where a conditional $\mathcal{B}el(B|A)$ is read as “if the agent were to learn A , she would believe B .” This interpretation was formalized in Conditional Doxastic Logic (**CDL**), originally introduced by Board [Boa04] in a game-theoretic setting, where it can be used to represent agents’ strategies, i.e. the choices they would make in response to possible moves of their opponents. **CDL** was later substantially developed by Baltag and Smets [BS08], who showed that it can serve as an epistemic formalism for belief revision in the style of AGM, internalizing the meta-logical revision operation as a conditional belief operator within the object language itself. They further extended **CDL** into a general framework for epistemic and doxastic reasoning, where conditional belief coexists with other notions such as knowledge, safe belief, and strong belief, and enriched it with dynamic modalities describing informational updates.

Semantically, **CDL** is defined using another variation of preferential semantics, where each agent is equipped both with an **S5**-accessibility relation and with a preference ordering over the accessible worlds. [GNOR18] observed that **CDL** can be viewed as a multi-agent extension of the counterfactual logic **VTA** (equivalent to Åqvist’s logic **G** from Chapter 3). They introduced an equivalent neighbourhood semantics closely related to Lewis’s sphere semantics for counterfactuals and developed proof-theoretic characterizations via labelled [GNOR18] and nested sequent calculi [GLO19]. These calculi maintain a close connection with the semantics and support countermodel reconstruction, but proof search and complexity analysis remain problematic: PSPACE-hardness was established in [GNOR18] and conjectured to be tight, but as far as we are aware, no explicit upper bound has been proposed so far.

Given the close connection between **CDL** and **G**, it is natural to expect that the methods developed in Chapter 4 can be extended to the multi-agent setting to obtain more streamlined analytic characterizations for CDL. On the semantic side, this would require integrating multi-agency into the Friedman–Halpern style small-model constructions, which can be done by branching not only over conditionals but also over agents. Unlike in the single-agent case of **G**, such branching can no longer be reduced to a single total preorder, plausibly reflecting the PSPACE-hardness of CDL. On the proof-theoretic side, one may extend the calculi from Sec. 4.8 by indexing all modalities, including the auxiliary $\boxtimes$ modality, with agents and treating them independently. Notably, while in the case of **PCL** and Åqvist’s logic (as in the original case of KLM logic **P**) $\boxtimes$ is introduced merely as an auxiliary semantic notion, in **CDL** it coincides with the modality of safe belief, already present in the logic. If successful, such an approach could be used, as in the case of Åqvist’s logics, to determine computational complexity of **CDL**, enable practically feasible automated deduction methods, and connect it with (multi-agent variants of) provability logics.

8.1.2 Grounded conditionals

The relational and preferential semantics considered throughout this thesis define modalities using abstract accessibility or preference relations characterized only by the properties they satisfy. An alternative approach, proposed in a series of works by Lorini [Lor18, Lor20] under the name of *grounded* modal logics, aims to “ground” such relations by deriving them from more elementary structures instead of taking them as primitive. Originally introduced in the epistemic setting, this framework equips each state not only with a propositional valuation, but also with a *belief base* for each agent, i.e. a set of formulas representing the agent’s explicit beliefs at the state; accessibility relations are then reconstructed as relations to states satisfying these formulas. This makes it possible to distinguish between explicit and implicit beliefs in the logical language via a modality Δ of explicit belief in addition to the standard modality $\Box$ of implicit belief, while also enabling more succinct representation of models without providing the accessibility relation explicitly. It also allows to naturally integrate different notions in one framework by defining them on the common underlying structure of belief bases. This grounded approach was later applied to a variety of epistemic notions, including distributed [HLP⁺20] and graded [LS21] belief, awareness [LS23], and only-knowing [dLLS23], and subsequently extended beyond epistemic logic to areas such as causal [Lor23, dLL24, AVLLR25] and deontic [CDL⁺25] reasoning and reasoning about cognitive attitudes [Lor22].

More recently, several works co-authored by the author of this thesis extended the grounded framework to conditional logics, now deriving preference relations from underlying bases of propositional facts. Grounded counterfactuals were defined over causal bases [AVLLR25], making it possible to represent notions such as actual causation directly in a counterfactual setting without relying on a primitive notion of intervention. Another approach [LR26] integrated counterfactuals with epistemic modalities by grounding them in belief bases, providing a unified framework for reasoning about beliefs, belief change, and counterfactual dependence. In the deontic setting, the logic of grounded conditional permissions was introduced [CDL⁺25], allowing to specify explicit “allowances” provided by system designer and derive permissions, obligations and prohibitions from them. While similar in spirit, these approaches differ technically in how the preference relation is computed from the underlying bases, adapting the construction to the particular application domain.

Analytic methods have so far been applied to grounded logics rather sparsely. Analytic tableau calculi were developed for the basic grounded epistemic logics and for several isolated extensions, but for grounded conditional logics no proof-theoretic characterizations are currently known, not even in the form of Hilbert-style axiomatizations. PSPACE-completeness was established for grounded conditional permissions by adapting Friedman-Halpern small-model constructions, but no comparable complexity results are known for the other two grounded conditional formalisms mentioned above.

The logic of grounded conditional permissions is particularly closely related to Åqvist’s framework: the induced preference relations satisfy the semantic conditions of Åqvist’s

logic $\mathbf{FCM}$, and conversely every $\mathbf{FCM}$ -model can be grounded by suitable permission bases, making grounded permissions a conservative extension of $\mathbf{FCM}$. This suggests that the $\mathbf{GL}$ -based proof-theoretic approach from Chapter 4 could be adapted to the grounded setting too, specifically by integrating Δ -modalities into the elimination rules for the auxiliary $\boxminus$ -modality to reflect the grounded definition of preference relations. Note that developing analytic proof calculi may be actually easier than finding axiomatisation for grounded logics, since reconstruction of an analytic countermodels for concrete entailment instances via proof search is more flexible than a standard approach of constructing a canonical model over maxconsistent set of arbitrary formulas. In particule, this was observed in [LR25] where some required properties, that are achieved in the canonical model by intricate model transformations, was integrated into inference rules directly. Once analytic calculi are available, corresponding axiomatizations can potentially be extracted from the inference rules themselves, although in this case only in the language extended with $\boxminus$ -modality.

8.2 Future perspectives

The results of this thesis demonstrate that analytic methods provide a powerful and unifying perspective on conditional reasoning, connecting proof theory, semantics, computational complexity, and automated deduction across several traditionally separate families of conditional logics. They also open a number of natural directions for future research.

Covering related logics (or families of logics). Beyond the five representative frameworks considered here, our approach is intended to extend systematically to substantially broader families of related logics. Examples of systems that we plan to investigate—or have already begun investigating—include:

- extensions of $\mathbf{PCL}$ with various frame properties [FH94, Gir19], including Lewis’s family of counterfactual logics [Lew73],
- further versions of I/O logics, such as aggregative I/O logics [PvdT14] or I/O logics with consistency checks [PvdT18], I/O logics for permissions [MvdT03, OPvdT23],
- variations of KLM logics obtained by removing some of the postulated (such as the version of logic $\mathbf{P}$ without cautious monotonicity posulate (*CMon*) in [ACPT25], capturing the notion of formula support to formalize computational trust) or adding new ones (such as logic $\mathbf{R}$ [GGOP09] that extends $\mathbf{R}$ with rational monotony principle),
- alternative closure operators such as lexicographic closure [Leh99], multipreference closure [GG21], and disjunctive rational closure [BV21].

Mechanisms of global non-monotonicity. Rational closure, studied in Chapter 7, adds global non-monotonicity extension to KLM-style entailment via minimization of ranked models. Globally non-monotonic extensions were also suggested for Input/Output logics, using quite different approaches: constrained Input/Output logics [MvdT01] employ maximally consistent subtheories similarly to the standard AGM approach to belief revision, while causal I/O logics admit notions of exact theories [Boc03] analogous to stable models in logic programming, which also yield globally non-monotonic entailment. Developing analytic characterizations for these frameworks would provide a basis for comparing different mechanisms of global non-monotonicity and understanding whether techniques developed in one formalism can be transferred to another. Such a comparison is particularly relevant because different mechanisms address different reasoning patterns: for example, rational closure captures defeasible inheritance and specificity, whereas constrained I/O logics naturally accommodate contrary-to-duty reasoning. Understanding these mechanisms within a common analytic framework may make it possible to combine their strengths and transfer techniques between different domains of conditional reasoning. For example, techniques underlying rational closure could potentially be adapted to normative reasoning to capture forms of specificity familiar from legal reasoning, such as legal principle “*lex specialis derogat legi generali*”, according to which more specific obligations should override more general ones.

Purely syntactic study of proposed calculi. Throughout the thesis, completeness and analyticity were established primarily via semantic methods, interpreting proof search as systematic construction of countermodels. This leaves open the possibility of studying the calculi from a more traditional Gentzen-style perspective, based on syntactic transformations such as cut elimination, as done, e.g. in [COP⁺23]. Note that unlike our semantical methods, cut elimination for calculi for **GL** and its relatives is notoriously challenging, especially when frame properties like linearity or modularity (corresponding to Åqvist logic **G**) are added. Such syntactical analysis could shed new light on structural properties of the considered logics, such as interpolation. Interpolation in particular may require alternative versions of the calculi for **PCL**, Åqvist’s logics, and rational closure, which avoid using modalities that do not belong to the language.

Small model constructions for weaker preferential models. Friedman-Halpern small model constructions, developed for **PCL** and its extensions, are applicable only for preferential models satisfying transitivity and smoothness. In Chapter 4, we developed an alternative small-model methodology for Åqvist’s logics, some of which lack these properties, while instead satisfying absoluteness — a property absent in **PCL**. It raises the question whether the LEGO-like approach can be extended to models without absoluteness and other generalizations of preferential models, such as typicality preferential models from Chapter 6, where individual worlds are replaced by multi-world states.

Optimal proof strategies and proof complexity. While for most of the developed calculi we proved that proof search matches the optimal complexity of the logic, it remains

open whether an optimal-complexity proof search strategy exists for the calculi developed for Åqvist’s logics in Chapter 4. A related issue arises for rational closure: there, we showed that a saturated hypersequent can be obtained within optimal complexity by exploiting the constructive computation of exceptionality ranks — essentially reproducing the known algorithmic characterization rather than yielding an intrinsically proof-theoretic procedure. An interesting open problem is therefore to identify a natural proof-search strategy of optimal complexity that does not appeal to exceptionality ranks. Closely related are questions of proof complexity. Since all considered logics are co-NP-hard, polynomial-size derivations cannot generally exist unless NP equals co-NP. Nevertheless, alternative representations of derivations — for instance, the ones that do not unfold purely classical fragments derivations with **LK** rules — may yield more succinct proof objects that could serve as practical certificates or explanations of entailment.

Provers for remaining systems. Efficient provers based on the presented characterizations were developed for Åqvist’s logics, original and causal I/O logics, and KLM logics. Comparable implementations are still missing for **PCL**, throughput I/O logics, and rational closure. Developing such tools is therefore a natural next step. More ambitiously, the substantial common structure uncovered in this thesis raises the possibility of a unified reasoning platform supporting multiple families of conditional logics within a single framework. Another important direction is a more extensive empirical evaluation of the existing and future provers, including systematic comparisons with alternative reasoning tools and a more detailed analysis of their practical scalability.

Overview of Generative AI Tools Used

The generative AI tool ChatGPT (version GPT-5.5, 2025) was used solely for text-editing, translation, and formatting assistance during the preparation of this thesis. All scientific content, technical details, and research contributions are entirely the author's own.

List of Figures

1.1	Map of the covered logics and obtained results	7
2.1	Propositional fragment of Gentzen's sequent calculus LK for classical logic.	19
2.2	Structural sequent rules	20
2.3	Sequent calculi for some normal modal logics.	20
2.4	Hypersequent calculus $\mathcal{H}_{\mathbf{S5}}$ for S5	21
2.5	Kleene-style hypersequent calculus $\mathcal{HK}_{\mathbf{S5}}$ for S5 . The addition in each premise w.r.t. the conclusion is highlighted.	23
2.6	Antisquent calculus <i>antiLK</i> for the classical non-validity	26
3.1	Rules of calculus $SC_{\mathbf{PCL}}$ for logic PCL	34
4.1	Preference-semantical characterizations for Åqvist's logics	50
4.2	LEGO small model constructions for Åqvist's logics	56
4.3	Kleene-style hypersequent rules for conditional obligations from [COP22]	71
5.1	Sequent calculi for the core I/O logics	91
5.2	Frame conditions of I/O models corresponding to I/O inference principles.	113
6.1	Saturating calculi for KLM logics	141
6.2	Derivation of $A \rightsquigarrow B, A \rightsquigarrow C \vdash A \wedge B \rightsquigarrow C$ in $\mathcal{SC}_{\mathbf{C}}$	143
6.3	Experimental evaluation for KLM provers from [CER24]	153
7.1	Rules of the basic calculus $\mathcal{HS}_{\mathcal{R}}$	161
7.2	Correspondence between non-entailment witnesses w.r.t. different characteri- zations of rational closure.	166

List of Tables

2.1	Axioms and frame conditions for some normal modal logics used in this thesis.	18
4.1	Finite-model characterizations of theoremhood in Åqvist's logics (with maximality as the notion of bestness).	68
5.1	Core I/O logics defined by I/O principles	90
5.2	Separation principles for the core I/O logics	92
5.3	Embeddings of $(A_1, X_1), \dots, (A_n, X_n) \vdash_{\mathcal{L}} (B, Y)$ into classical propositional logic for every core I/O logic $\mathcal{L}$	118
5.4	Embeddings of the core I/O logics into normal modal logics.	120
6.1	The axiomatic definition of KLM logics	137
6.2	Conditions on preference models for KLM logics	138

List of Algorithms

5.1	Algorithm for deciding if G entails $(\mathcal{B}, y)$ in an I/O logic $\mathcal{L}$ without (OR) for the literal fragment	127
-----	---	-----

Bibliography

- [AA00] Ofer Arieli and Arnon Avron. General patterns for nonmonotonic reasoning: From basic entailments to plausible relations. *Logic Journal of the IGPL*, 8(2):119–148, 2000.
- [ACPT25] Alessandro Aldini, Agata Ciabattone, Dominik Pichler, and Mirko Tagliaferri. Support + belief = decision trust. In *Structural Information and Communication Complexity: 32nd International Colloquium, SIROCCO 2025, Delphi, Greece, June 2–4, 2025, Proceedings*, page 3–22, Berlin, Heidelberg, 2025. Springer-Verlag.
- [Ada75] Ernest W. Adams. *The Logic of Conditionals*. D. Reidel, 1975.
- [AOP16] Régis Alenda, Nicola Olivetti, and Gian Luca Pozzato. Nested sequent calculi for normal conditional logics. *J. Log. Comput.*, 26(1):7–50, 2016.
- [Åqv84] Lennart Åqvist. Deontic logic. In D. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic: Volume II*, pages 605–714. Springer, Dordrecht, 1984.
- [Åqv02] Lennart Åqvist. Deontic logic. In D. M. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic: Volume 8*, pages 147–264. Springer Netherlands, Dordrecht, 2002.
- [AvBRS24] Ofer Arieli, Kees van Berkel, Badran Raddaoui, and Christian Straßer. Deontic reasoning based on inconsistency measures. In Pierre Marquis, Magdalena Ortiz, and Maurice Pagnucco, editors, *Proceedings of the 21st International Conference on Principles of Knowledge Representation and Reasoning, KR 2024*, 2024.
- [AvBS24] Ofer Arieli, Kees van Berkel, and Christian Straßer. Defeasible normative reasoning: A proof-theoretic integration of logical argumentation. In Michael J. Wooldridge, Jennifer G. Dy, and Sriraam Natarajan, editors, *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence*,

EAAI 2014, February 20-27, 2024, Vancouver, Canada, pages 10450–10458. AAAI Press, 2024.

- [AVLLR25] Carlos Aguilera-Ventura, Xuhui Liu, Emiliano Lorini, and Dmitry Rozplokhas. A non-interventionist approach to causal reasoning based on lewisian counterfactuals. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2025)*, pages 4301–4310. ijcai.org, 2025.
- [Avr96a] A. Avron. The method of hypersequents in the proof theory of propositional non-classical logics. In *Logic: from foundations to applications*, pages 1–32. OUP, New York, 1996.
- [Avr96b] Arnon Avron. The method of hypersequents in the proof theory of propositional non-classical logics. In Wilfrid Hodges, editor, *Logic: From Foundations to Applications*, pages 1–32. Oxford University Press, New York, 1996.
- [BCM⁺21] Katarina Britz, Giovanni Casini, Thomas Meyer, Kody Moodley, Uli Sattler, and Ivan Varzinczak. Principles of KLM-style defeasible description logics. *ACM Trans. Comput. Log.*, 22(1):1:1–1:46, 2021.
- [Bel82] Nuel D. Belnap. Display logic. *Journal of Philosophical Logic*, 11(4):375–417, 1982.
- [Bet55] Evert Willem Beth. *Semantic Entailment and Formal Derivability*. North-Holland, Amsterdam, 1955.
- [BFMP19] Christoph Benzmüller, Ali Farjami, Paul Meder, and Xavier Parent. I/O logic in HOL. *FLAP*, 6(5):715–732, 2019.
- [BHM09] Katarina Britz, Johannes Heidema, and Thomas Meyer. Semantics for dual preferential entailment. *Journal of Philosophical Logic*, 38(4):423–451, 2009.
- [BL15] Alexander Bochman and Vladimir Lifschitz. Pearl’s causality in a logical setting. In Blai Bonet and Sven Koenig, editors, *Proceedings of AAAI 2015*, pages 1446–1452. AAAI Press, 2015.
- [BO97] Piero Andrea Bonatti and Nicola Olivetti. A sequent calculus for skeptical default logic. In *Automated Reasoning with Analytic Tableaux and Related Methods: 6th International Workshop, TABLEAUX ’97*, volume 1227 of *Lecture Notes in Artificial Intelligence*, pages 107–121, Berlin, Heidelberg, 1997. Springer.
- [BO98] Piero A. Bonatti and Nicola Olivetti. A sequent calculus for circumscription. In *Computer Science Logic: 11th International Workshop, CSL ’97, Selected Papers*, volume 1414 of *Lecture Notes in Computer Science*, pages 98–114, Berlin, Heidelberg, 1998. Springer.

- [BO02] Piero Andrea Bonatti and Nicola Olivetti. Sequent calculi for propositional nonmonotonic logics. *ACM Trans. Comput. Logic*, 3(2):226–278, April 2002.
- [Boa04] Oliver Board. Dynamic interactive epistemology. *Games and Economic Behavior*, 49(1):49–80, 2004.
- [Boc03] Alexander Bochman. A logic for causal reasoning. In Georg Gottlob and Toby Walsh, editors, *Proceedings of IJCAI-03*, pages 141–146. Morgan Kaufmann, 2003.
- [Boc04] Alexander Bochman. A causal approach to nonmonotonic reasoning. *Artif. Intell.*, 160(1-2):105–143, 2004.
- [Boc06] Alexander Bochman. Two paradigms of nonmonotonic reasoning. In *International Symposium on Artificial Intelligence and Mathematics, AI&Math 2006, Fort Lauderdale, Florida, USA, January 4-6, 2006*, 2006.
- [Boc07] Alexander Bochman. A causal theory of abduction. *J. Log. Comput.*, 17(5):851–869, 2007.
- [Boc18] Alexander Bochman. Actual causality in a logical setting. In Jérôme Lang, editor, *Proceedings of IJCAI 2018*, pages 1730–1736. ijcai.org, 2018.
- [Boc21] Alexander Bochman. *A Logical Theory of Causality*. MIT Press, 2021.
- [Boc23] Alexander Bochman. An inferential theory of causal reasoning. In Natasha Alechina, Andreas Herzig, and Fei Liang, editors, *Logic, Rationality, and Interaction - 9th International Workshop, LORI 2023, Jinan, China, October 26-29, 2023, Proceedings*, Lecture Notes in Computer Science, pages 1–16. Springer, 2023.
- [Bon93] Piero Bonatti. A gentzen system for non-theorems. Technical Report CD-TR 93/52, Christian Doppler Labor für Expertensysteme, Technische Universität Wien, Vienna, Austria, 1993.
- [Bon96] Piero Andrea Bonatti. Sequent calculi for default and autoepistemic logics. In *Automated Reasoning with Analytic Tableaux and Related Methods: 5th International Workshop, TABLEAUX '96*, volume 1071 of *Lecture Notes in Artificial Intelligence*, pages 127–142, Berlin, Heidelberg, 1996. Springer.
- [Bon19] Piero Andrea Bonatti. Rational closure for all description logics. *Artificial Intelligence*, 274:197–223, 2019.
- [Bou90] Craig Boutilier. Conditional logics of normality as modal systems. In Howard E. Shrobe, Thomas G. Dietterich, and William R. Swartout, editors, *Proceedings of the 8th National Conference on Artificial Intelligence. Boston, Massachusetts, USA, July 29 - August 3, 1990, 2 Volumes*, pages 594–599. AAAI Press / The MIT Press, 1990.

- [BP98] Richard Booth and Jeffrey Paris. A note on the rational closure of knowledge bases with both positive and negative knowledge. *Journal of Logic, Language and Information*, 7(2):165–190, 1998.
- [BPvdT09] Guido Boella, Gabriella Pigozzi, and Leendert van der Torre. Normative framework for normative system change. In Carles Sierra, Cristiano Castelfranchi, Keith S. Decker, and Jaime Simão Sichman, editors, *8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009), Budapest, Hungary, May 10-15, 2009, Volume 1*, pages 169–176. IFAAMAS, 2009.
- [BPvdT20] Christoph Benzmüller, Xavier Parent, and Leendert van der Torre. Designing normative theories for ethical and legal reasoning: Logikey framework, methodology, and tool support. *Artif. Intell.*, 287:103348, 2020.
- [Brü09] Kai Brunnler. Deep sequent systems for modal logic. *Archive for Mathematical Logic*, 48(6):551–577, 2009.
- [BS08] Alexandru Baltag and Sonja Smets. A qualitative theory of dynamic interactive belief revision. *Texts in Logic and Games*, 3:9–58, 2008.
- [Bur81] John P. Burgess. Quick completeness proofs for some logics of conditionals. *Notre Dame Journal of Formal Logic*, 22(1):76–84, 1981.
- [BV18] Katarina Britz and Ivan Varzinczak. From klm-style conditionals to defeasible modalities, and back. *Journal of Applied Non-Classical Logics*, 28(1):92–121, 2018.
- [BV21] Richard Booth and Ivan Varzinczak. Conditional inference under disjunctive rationality. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pages 6227–6234. AAAI Press, 2021.
- [BvdTV07] Guido Boella, Leon van der Torre, and Harko Verhagen, editors. *Normative Multi-Agent Systems*. Dagstuhl Seminar Proceedings, 2007.
- [CDL⁺25] Agata Ciabattoni, Josephine Femke Dik, Emiliano Lorini, Dominik Pichler, and Dmitry Rozplokhas. From explicit allowances to defeasible deontic operators: A modal view. In Corina Dima, Angelo Ferrando, and Vadim Malvone, editors, *PRIMA 2025: Principles and Practice of Multi-Agent Systems*, volume 16366 of *Lecture Notes in Computer Science*, pages 585–603, Cham, 2025. Springer.
- [CER24] Agata Ciabattoni, Clemens Eisenhofer, and Dmitry Rozplokhas. Strongly analytic calculi for KLM logics with smt-based prover. In Pierre Marquis,

- Magdalena Ortiz, and Maurice Pagnucco, editors, *Proceedings of the 21st International Conference on Principles of Knowledge Representation and Reasoning, KR 2024, Hanoi, Vietnam. November 2-8, 2024*, 2024.
- [CGT08] Agata Ciabattoni, Nikolaos Galatos, and Kazushige Terui. From axioms to analytic rules in nonclassical logics. In *Proceedings of the Twenty-Third Annual IEEE Symposium on Logic in Computer Science, LICS 2008, 24-27 June 2008, Pittsburgh, PA, USA*, pages 229–240. IEEE Computer Society, 2008.
- [Chi63] Roderick Chisholm. Contrary-to-duty imperatives and deontic logic. *Analysis*, 24:33–36, 1963.
- [COP22] Agata Ciabattoni, Nicola Olivetti, and Xavier Parent. Dyadic obligations: Proofs and countermodels via hypersequents. In Reyhan Aydogan, Natalia Criado, Jérôme Lang, Víctor Sánchez-Anguix, and Marc Serramia, editors, *PRIMA 2022: Principles and Practice of Multi-Agent Systems - 24th International Conference, Valencia, Spain, November 16-18, 2022, Proceedings*, volume 13753 of *Lecture Notes in Computer Science*, pages 54–71. Springer, 2022.
- [COP⁺23] Agata Ciabattoni, Nicola Olivetti, Xavier Parent, Revantha Ramanayake, and Dmitry Rozplokhas. Analytic proof theory for Åqvist’s system F. In Juliano Maranhão, Clayton Peterson, Christian Straßer, and Leendert van der Torre, editors, *Deontic Logic and Normative Systems - 16th International Conference, DEON 2023, Trois-Rivières, QC, Canada, July 5-7, 2023*, pages 79–98. College Publications, 2023.
- [CPS21] Agata Ciabattoni, Xavier Parent, and Giovanni Sartor. A kelsenian deontic logic. In Schweighofer Erich, editor, *Proceedings of JURIX 2021*, volume 346 of *Frontiers in Artificial Intelligence and Applications*, pages 141–150. IOS Press, 2021.
- [CS10] Giovanni Casini and Umberto Straccia. Rational closure for defeasible description logics. In Tomi Janhunen and Ilkka Niemelä, editors, *Logics in Artificial Intelligence*, pages 77–90, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
- [CT24] Agata Ciabattoni and Matteo Tesi. Sequents vs hypersequents for Åqvist systems. In *The proceedings of International Joint Conference on Automated Reasoning (IJCAR)*, 2024.
- [DEGV01] Evgeny Dantsin, Thomas Eiter, Georg Gottlob, and Andrei Voronkov. Complexity and expressive power of logic programming. *ACM Comput. Surv.*, 33(3):374–425, 2001.

- [DFM⁺22] Andrea De Domenico, Ali Farjami, Krishna Manoorkar, Alessandra Palmigiano, Mattia Panettiere, and Xiaolong Wang. Subordination algebras as semantic environment of input/output logic. In Agata Ciabattoni, Elaine Pimentel, and Ruy J. G. B. de Queiroz, editors, *Logic, Language, Information, and Computation - 28th International Workshop, WoLLIC 2022, Iași, Romania, September 20-23, 2022, Proceedings*, Lecture Notes in Computer Science, pages 326–343. Springer, 2022.
- [dLL24] Tiago de Lima and Emiliano Lorini. Model checking causality. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI 2024)*, pages 3324–3332. International Joint Conferences on Artificial Intelligence Organization, 2024.
- [dLLS23] Tiago de Lima, Emiliano Lorini, and François Schwarzentruher. Base-based model checking for multi-agent only believing. In *Logics in Artificial Intelligence: 18th European Conference, JELIA 2023, Dresden, Germany, September 20–22, 2023, Proceedings*, volume 14281 of *Lecture Notes in Computer Science*, pages 437–445, Cham, 2023. Springer.
- [Ein24] Marit Einböck. Encoding of the entailment problem for input/output logics in classical propositional logic, 6 2024.
- [EL00] Thomas Eiter and Thomas Lukasiewicz. Default reasoning from conditional knowledge bases: Complexity and tractable cases. *Artificial Intelligence*, 124(2):169–241, 2000.
- [EL02] Thomas Eiter and Thomas Lukasiewicz. Complexity results for structure-based causality. *Artificial Intelligence*, 142(1):53–89, 2002.
- [FH94] Nir Friedman and Joseph Y. Halpern. On the complexity of conditional logics. In Jon Doyle, Erik Sandewall, and Pietro Torasso, editors, *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning (KR'94). Bonn, Germany, May 24-27, 1994*, pages 202–213. Morgan Kaufmann, 1994.
- [Gab85] Dov M. Gabbay. Theoretical foundations for non-monotonic reasoning in expert systems. In Krzysztof R. Apt, editor, *Logics and Models of Concurrent Systems*, pages 439–457, Berlin, Heidelberg, 1985. Springer Berlin Heidelberg.
- [Gab96] Dov M. Gabbay. *Labelled Deductive Systems*. Oxford University Press, Oxford, 1996.
- [Gen34] G. Gentzen. Untersuchungen über das logische Schließen. 39:176–210, 405–431, 1934. English translation in: *American Philosophical Quarterly* 1 (1964), 288–306 and *American Philosophical Quarterly* 2 (1965), 204–218, as well as in: *The Collected Papers of Gerhard Gentzen*, (ed. M.E. Szabo), Amsterdam, North Holland (1969), 68–131.

- [Gen35] Gerhard Gentzen. Untersuchungen über das logische Schließen. 39:176–210, 405–431, 1934/35. English translation in: *American Philosophical Quarterly* 1 (1964), 288–306 and *American Philosophical Quarterly* 2 (1965), 204–218, as well as in: *The Collected Papers of Gerhard Gentzen*, (ed. M.E. Szabo), Amsterdam, North Holland (1969), 68–131.
- [GG21] Laura Giordano and Valentina Gliozzi. A reconstruction of multipreference closure. *Artificial Intelligence*, 290:103398, 2021.
- [GGOP07] Laura Giordano, Valentina Gliozzi, Nicola Olivetti, and Gian Luca Pozzato. An implementation of a free-variable tableaux for KLM preferential logic P of nonmonotonic reasoning: The theorem prover FreeP 1.0. In Roberto Basili and Maria Teresa Pazienza, editors, *AI*IA*, volume 4733 of *LNCS*, pages 84–96. Springer, 2007.
- [GGOP09] Laura Giordano, Valentina Gliozzi, Nicola Olivetti, and Gian Luca Pozzato. Analytic tableaux calculi for KLM logics of nonmonotonic reasoning. *ACM Trans. Comput. Log.*, 10(3):18:1–18:47, 2009.
- [GGOP15] Laura Giordano, Valentina Gliozzi, Nicola Olivetti, and Gian Luca Pozzato. Semantic characterization of rational closure: From propositional logic to description logics. *Artificial Intelligence*, 226:1–33, 2015.
- [GGOS09] Laura Giordano, Valentina Gliozzi, Nicola Olivetti, and Camilla Schwind. Tableau calculus for preference-based conditional logics: PCL and its extensions. *ACM Trans. Comput. Log.*, 10(3), April 2009.
- [GGP07] Laura Giordano, Valentina Gliozzi, and Gian Luca Pozzato. KLMLean 2.0: A theorem prover for KLM logics of nonmonotonic reasoning. In Nicola Olivetti, editor, *TABLEAUX*, pages 238–244, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.
- [GHP⁺13] Dov M. Gabbay, John F. Horty, Xavier Parent, Ron van der Meyden, and Leendert van der Torre. *Handbook of deontic logic and normative systems*. 2013.
- [Gin85] Matthew L. Ginsberg. Counterfactuals. In *Proceedings of the 9th International Joint Conference on Artificial Intelligence (IJCAI’85)*, pages 80–86, 1985.
- [Gir19] Marianna Girlando. *On the Proof Theory of Conditional Logics*. Theses, Aix-Marseille Universite; Helsinki University, February 2019.
- [GLO19] Marianna Girlando, Björn Lellmann, and Nicola Olivetti. Nested sequents for the logic of conditional belief. In Francesco Calimeri, Nicola Leone, and Marco Manna, editors, *Logics in Artificial Intelligence: 16th European Conference, JELIA 2019, Rende, Italy, May 7–11, 2019, Proceedings*, volume

- 11468 of *Lecture Notes in Artificial Intelligence*, pages 709–725, Cham, 2019. Springer.
- [GLOP16] Marianna Girlando, Björn Lellmann, Nicola Olivetti, and Gian Luca Pozzato. Standard sequent calculi for lewis’ logics of counterfactuals. In Loizos Michael and Antonis C. Kakas, editors, *Logics in Artificial Intelligence: 15th European Conference, JELIA 2016, Larnaca, Cyprus, November 9–11, 2016, Proceedings*, volume 10021 of *Lecture Notes in Computer Science*, pages 272–287, Cham, 2016. Springer.
 - [GNO21] Marianna Girlando, Sara Negri, and Nicola Olivetti. Uniform labelled calculi for preferential conditional logics based on neighbourhood semantics. *J. Log. Comput.*, 31(3):947–997, 2021.
 - [GNOR18] Marianna Girlando, Sara Negri, Nicola Olivetti, and Vincent Risch. Conditional beliefs: From neighbourhood semantics to sequent calculus. *Review of Symbolic Logic*, 11(4):736–779, 2018.
 - [Goo47] Nelson Goodman. The problem of counterfactual conditionals. *The Journal of Philosophy*, 44(5):113–128, 1947.
 - [GOP13] Laura Giordano, Nicola Olivetti, and Gian Luca Pozzato. On rational closure in description logics of typicality. 05 2013.
 - [Han69] Bengt Hansson. An analysis of some deontic logics. *Noûs*, 3(4):373–398, 1969.
 - [HGY22] Sven Ove Hansson and Till Grüne-Yanoff. Preferences. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2022 edition, 2022.
 - [Hin62] Jaakko Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca, NY, 1962.
 - [HLP⁺20] A. Herzig, E. Lorini, E. Perrotin, F. Romero, and F. Schwarzentruher. A logic of explicit and implicit distributed belief. In *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI 2020)*, volume 325 of *Frontiers in Artificial Intelligence and Applications*, pages 753–760. IOS Press, 2020.
 - [Hop25] Brandon Hopkins. Input/output logic via neighborhood semantics. In Kees van Berkel, Agata Ciabattoni, and John F. Horty, editors, *Deontic Logic and Normative Systems - 17th International Conference, DEON 2025, Vienna, Austria, July 1-3, 2025*, pages 179–198. College Publications, 2025.
 - [HP05] Joseph Y. Halpern and Judea Pearl. Causes and explanations: a structural-model approach. Part I: Causes. *British Journal for Philosophy of Science*, 56(4):843–887, 2005.

- [J37] Jürgen Jörgensen. Imperatives and logic. *Erkenntnis*, 7(1):288–296, 1937.
- [KI19] Gabriele Kern-Isberner. Conditionals in nonmonotonic reasoning and belief revision: Considering conditionals as agents. In Angelika Kimmig, Bernhard Nebel, and Matthias Thimm, editors, *Conditionals and Conditional Information: Sources and Use*, volume 313 of *Frontiers in Artificial Intelligence and Applications*, pages 1–18. IOS Press, 2019.
- [KL18] Roman Kuznets and Björn Lellmann. G3-style sequent calculi for modal logics. *IfCoLog Journal of Logics and their Applications*, 5(3):463–514, 2018.
- [KLM90] Sarit Kraus, Daniel Lehmann, and Menachem Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artif. Intell.*, 44(1-2):167–207, 1990.
- [Köl25] Christian Köll. Smt-based automated reasoning for åqvist’s deontic logics. Master’s thesis, TU Wien, 2025.
- [Kri66] Saul A. Kripke. Semantical analysis of modal logic i. normal modal propositional calculi. *Journal of Symbolic Logic*, 31(1):120–122, 1966.
- [KS91] Hirofumi Katsuno and Ken Satoh. A unified view of consequence relation, belief revision and conditional logic. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI ’91)*, pages 406–412. Morgan Kaufmann, 1991.
- [Kur13] Hidenori Kurokawa. Hypersequent calculi for modal logics extending S4. In *New Frontiers in Artificial Intelligence*, volume 8417 of *LNCS*, pages 51–68. Springer, 2013.
- [Lad77] Richard E. Ladner. The computational complexity of provability in systems of modal propositional logic. *SIAM Journal on Computing*, 6(3):467–480, 1977.
- [Lah13] Ori Lahav. From frame properties to hypersequent rules in modal logics. In *28th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2013, New Orleans, LA, USA, June 25-28, 2013*, pages 408–417. IEEE Computer Society, 2013.
- [LCG⁺23] Tim S. Lyon, Agata Ciabattoni, Didier Galmiche, Dominique Larchey-Wendling, Daniel Méry, Nicola Olivetti, and Revantha Ramanayake. Internal and external calculi: Ordering the jungle without being lost in translations. *CoRR*, abs/2312.03426, 2023.
- [Leh99] Daniel Lehmann. Another perspective on default reasoning. *Annals of Mathematics and Artificial Intelligence*, 15, 11 1999.

- [Lel21] Björn Lellmann. From input/output logics to conditional logics via sequents - with provers. In Anupam Das and Sara Negri, editors, *TABLEAUX*, volume 12842 of *LNCS*, pages 147–164. Springer, 2021.
- [Lew73] David Lewis. *Counterfactuals*. Blackwell, Oxford, 1973.
- [Lew79] David Lewis. Counterfactual dependence and time’s arrow. *Noûs*, 13(4):455–476, 1979.
- [Lew86] David Lewis. Causation. In *Philosophical Papers, Volume II*, pages 159–213. Oxford University Press, Oxford, 1986.
- [LL32] Clarence Irving Lewis and Cooper Harold Langford. *Symbolic Logic*. Century, 1932.
- [LM92] Daniel Lehmann and Menachem Magidor. What does a conditional knowledge base entail? *Artif. Intell.*, 55(1):1–60, 1992.
- [Löb55] M. H. Löb. Solution of a problem of leon henkin. *The Journal of Symbolic Logic*, 20(2):115–118, 1955.
- [Lor18] Emiliano Lorini. In praise of belief bases: Doing epistemic logic without possible worlds. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*, pages 1915–1922. AAAI Press, 2018.
- [Lor20] Emiliano Lorini. Rethinking epistemic logic with belief bases. *Artificial Intelligence*, 282, 2020.
- [Lor22] Emiliano Lorini. A computationally grounded framework for cognitive attitudes. *Artificial Intelligence*, 302:103602, 2022.
- [Lor23] Emiliano Lorini. A rule-based modal view of causal reasoning. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-23)*, pages 3286–3295. International Joint Conferences on Artificial Intelligence Organization, 2023.
- [LR25] Emiliano Lorini and Dmitry Rozplokhas. Graded distributed belief. In *Proceedings of the 19th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2025)*, pages 203–215, 2025.
- [LR26] Emiliano Lorini and Dmitry Rozplokhas. Merging belief and counterfactuals. In *Proceedings of the Thirty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-ECAI 2026)*. International Joint Conferences on Artificial Intelligence Organization, 2026. To appear.
- [LS77] E. J. Lemmon and Dana Scott. *An Introduction to Modal Logic*. Number 11 in American Philosophical Quarterly Monograph Series. Basil Blackwell, Oxford, 1977.

- [LS21] Emiliano Lorini and François Schwarzentruher. A computationally grounded logic of graded belief. In *Proceedings of the 17th European Conference on Logics in Artificial Intelligence (JELIA 2021)*, volume 12678 of *LNCS*, pages 245–261. Springer, 2021.
- [LS23] Emiliano Lorini and Pengfei Song. A computationally grounded logic of awareness. *Journal of Logic and Computation*, 33(6):1463–1496, 2023.
- [Łuk64] Jan Łukasiewicz. *Elements of Mathematical Logic*. Pergamon Press, 1964.
- [Mak89] David Makinson. General theory of cumulative inference. In Michael Reinfrank, Johan de Kleer, Matthew L. Ginsberg, and Erik Sandewall, editors, *Non-Monotonic Reasoning*, volume 346 of *Lecture Notes in Artificial Intelligence*, pages 1–18. Springer, Berlin, Heidelberg, 1989.
- [Mak94] David Makinson. General patterns in nonmonotonic reasoning. In Dov M. Gabbay, C. J. Hogger, J. A. Robinson, and Donald Nute, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3, pages 35–110. Oxford University Press, 1994.
- [McC80] John McCarthy. Circumscription — a form of non-monotonic reasoning. *Artificial Intelligence*, 13(1–2):27–39, 1980.
- [MOG08] George Metcalfe, Nicola Olivetti, and Dov M. Gabbay. *Proof Theory for Fuzzy Logics*, volume 36 of *Applied Logic Series*. Springer, Dordrecht, 2008.
- [Moo85] Robert C. Moore. Semantical considerations on nonmonotonic logic. *Artificial Intelligence*, 25(1):75–94, 1985.
- [MT97] Norman McCain and Hudson Turner. Causal theories of action and change. In Benjamin Kuipers and Bonnie L. Webber, editors, *Proceedings of AAAI 97*, pages 460–465. AAAI Press / The MIT Press, 1997.
- [MvdT00] David Makinson and Leendert van der Torre. Input/output logics. *J. Philos. Log.*, 29(4):383–408, 2000.
- [MvdT01] David Makinson and Leendert van der Torre. Constraints for input/output logics. *J. Philos. Log.*, 30(2):155–185, 2001.
- [MvdT03] David Makinson and Leendert van der Torre. Permission from an input/output perspective. *Journal of Philosophical Logic*, 32:391–416, 2003.
- [Neg05] Sara Negri. Proof analysis in modal logic. *Journal of Philosophical Logic*, 34(5–6):507–544, 2005.
- [NO15] Sara Negri and Nicola Olivetti. A sequent calculus for preferential conditional logic based on neighbourhood semantics. *Proceedings of Tableaux*, 9323:115–134, 2015.

- [NP18] Cláudia Nalon and Dirk Pattinson. A resolution-based calculus for preferential logics. In Didier Galmiche, Stephan Schulz, and Roberto Sebastiani, editors, *IJCAR 2018, Proceedings*, volume 10900 of *Lecture Notes in Computer Science*, pages 498–515. Springer, 2018.
- [NvP11] Sara Negri and Jan von Plato. *Proof Analysis: A Contribution to Hilbert’s Last Problem*. Cambridge University Press, Cambridge, 2011.
- [OP14] Nicola Olivetti and Gian Luca Pozzato. NESCOND: an implementation of nested sequent calculi for conditional logics. In Stéphane Demri, Deepak Kapur, and Christoph Weidenbach, editors, *IJCAR*, volume 8562 of *LNCS*, pages 511–518. Springer, 2014.
- [OPvdT23] Maya Olszewski, Xavier Parent, and Leendert van der Torre. Permissive and regulative norms in deontic logic. *Journal of Logic and Computation*, 34:728–763, 2023.
- [Par14] Xavier Parent. Maximality vs. optimality in dyadic deontic logic. *J. Philos. Log.*, 43(6):1101–1128, 2014.
- [Par21] Xavier Parent. Preference semantics for Hansson-type dyadic deontic logic: a survey of results. In D. Gabbay, J. Horty, X. Parent, L. van der Torre, and R. van der Meyden, editors, *Handbook of Deontic Logic and Normative Systems (vol. 2)*, pages 7–70. College Publications, London, 2021.
- [Par26] Xavier Parent. Defeasible conditional obligation in a two-tier preference-based semantics. In *Proceedings of the 21st International Conference on Principles of Knowledge Representation and Reasoning, KR 2026 (to appear)*, 2026.
- [PB23] Xavier Parent and Christoph Benzmüller. Automated verification of deontic correspondences in isabelle/hol – first results. In Christoph Benzmüller and Jens Otten, editors, *ARQNL 2022: Automated Reasoning in Quantified Non-Classical Logics. Proceedings of the 4th International Workshop on Automated Reasoning in Quantified Non-Classical Logics (ARQNL 2022) affiliated with the 11th International Joint Conference on Automated Reasoning (IJCAR 2022)*, volume 3326 of *CEUR Workshop Proceedings*, pages 92–108. CEUR-WS.org, 2023.
- [Pea90] Judea Pearl. System z: a natural ordering of defaults with tractable applications to nonmonotonic reasoning. In *Proceedings of the 3rd Conference on Theoretical Aspects of Reasoning about Knowledge, TARK ’90*, page 121–135, San Francisco, CA, USA, 1990. Morgan Kaufmann Publishers Inc.
- [Pog11] Francesca Poggioli. *Gentzen Calculi for Modal Propositional Logic*, volume 32 of *Trends in Logic*. Springer, 2011.

- [Pot83] Garrel Pottinger. Uniform, cut-free formulations of T, S4 and S5 (abstract). *J. of Symbolic Logic*, 48(3):900, 1983.
- [PS14] Frederik Van De Putte and Christian Straßer. Preferential semantics using non-smooth preference relations. *Journal of Philosophical Logic*, 43(5):903–942, 2014.
- [PvdT14] Xavier Parent and Leon van der Torre. Aggregative deontic detachment for normative reasoning. In Chitta Baral, Giuseppe De Giacomo, and Thomas Eiter, editors, *Proceedings of KR 2014*. AAAI Press, 2014.
- [PvdT18] Xavier Parent and Leendert van der Torre. I/O logics with a consistency check. In Jan M. Broersen, Cleo Condoravdi, Nair Shyam, and Gabriella Pigozzi, editors, *Proceedings of DEON 2018*, pages 285–299. College Publications, 2018.
- [RBP⁺20] Livio Robaldo, Cesare Bartolini, Monica Palmirani, Arianna Rossi, Michele Martoni, and Gabriele Lenzini. Formalizing GDPR provisions in reified I/O logic: The DAPRECO knowledge base. *J. Log. Lang. Inf.*, 29(4):401–449, 2020.
- [Rei80] Raymond Reiter. A logic for default reasoning. *Artificial Intelligence*, 13(1–2):81–132, 1980.
- [SBP16] Christian Straßer, Mathieu Beirlaen, and Frederik Van De Putte. Adaptive logic characterizations of input/output logic. *Studia Logica*, 104(5):869–916, 2016.
- [Seg71] Krister Segerberg. *An Essay in Classical Modal Logic*, volume 13 of *Filosofiska Studier*. Filosofiska föreningen och Filosofiska institutionen vid Uppsala universitet, Uppsala, 1971.
- [Sho87] Yoav Shoham. A semantical approach to non-monotonic logics. In *LICS*, pages 275–279. IEEE Computer Society Press, 1987.
- [Smu68] Raymond M. Smullyan. *First-Order Logic*. Springer, Berlin, 1968.
- [Sol76] Robert M. Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics*, 25(3):287–304, 1976.
- [SPH10] Lutz Schröder, Dirk Pattinson, and Daniel Hausmann. Optimal tableaux for conditional logics with cautious monotonicity. In Helder Coelho, Rudi Studer, and Michael J. Wooldridge, editors, *ECAI 2010, Proceedings*, volume 215 of *Frontiers in Artificial Intelligence and Applications*, pages 707–712. IOS Press, 2010.
- [SR17] Xin Sun and Livio Robaldo. On the complexity of input/output logic. *J. Appl. Log.*, 25:69–88, 2017.

- [Sta68] Robert Stalnaker. A theory of conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory*, pages 98–112. Blackwell, 1968.
- [Sta94] Robert Stalnaker. What is a nonmonotonic consequence relation? *Fundamenta Informaticae*, 21(1–2):7–21, 1994.
- [Ste25] Alexander Steen. A reduction of input/output logics to SAT. *CoRR*, abs/2508.16242, 2025.
- [Sto15] Audun Stolpe. A concept approach to input/output logic. *J. Appl. Log.*, 13(3):239–258, 2015.
- [Str09] Christian Straßer. An adaptive logic for rational closure. In Walter Carnielli, Marcello E. Coniglio, and Itala M. Loffredo D’Ottaviano, editors, *The Many Sides of Logic*, volume 21 of *Studies in Logic*, pages 47–67. College Publications, London, UK, 2009.
- [Sun14] Xin Sun. How to build input/output logic. In Nils Bulling, Leendert van der Torre, Serena Villata, Wojtek Jamroga, and Wamberto Weber Vasconcelos, editors, *Computational Logic in Multi-Agent Systems - 15th International Workshop, CLIMA XV, Prague, Czech Republic, August 18-19, 2014. Proceedings*, Lecture Notes in Computer Science, pages 123–137. Springer, 2014.
- [SV82] Giovanni Sambin and Silvio Valentini. The modal logic of provability. the sequential approach. *Journal of Philosophical Logic*, 11(3):311–342, 1982.
- [Val83] Silvio Valentini. The modal logic of provability: cut-elimination. *J. Philos. Logic*, 12(4):471–476, 1983.
- [vBS22] Kees van Berkel and Christian Straßer. Reasoning with and about norms in logical argumentation. In Francesca Toni, Sylwia Polberg, Richard Booth, Martin Caminada, and Hiroyuki Kido, editors, *Computational Models of Argument - Proceedings of COMMA 2022, Cardiff, Wales, UK, 14-16 September 2022*, Frontiers in Artificial Intelligence and Applications, pages 332–343. IOS Press, 2022.
- [vBS25] Kees van Berkel and Christian Straßer. Reasoning with and about normative conflicts. In Kees van Berkel, Agata Ciabattoni, and John F. Horty, editors, *Deontic Logic and Normative Systems - 17th International Conference, DEON 2025, Vienna, Austria, July 1-3, 2025*, pages 21–40. College Publications, 2025.
- [vdTP13] Leendert van der Torre and Xavier Parent. Input/output logics. In Dov M. Gabbay, John F. Horty, Xavier Parent, Ron van der Meyden, and Leendert van der Torre, editors, *Handbook of deontic logic and normative systems*, pages 499–544, 2013.

- [vW51] Georg Henrik von Wright. Deontic logic. *Mind*, 60(237):1–15, 1951.
- [Wan98] Heinrich Wansing. *Displaying Modal Logic*, volume 3 of *Trends in Logic*. Kluwer Academic Publishers, Dordrecht, 1998.
- [Woo03] James Woodward. *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, Oxford, 2003.