Tableaux and Hypersequents for Modal and Justification Logics

by

HIDENORI KUROKAWA

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2012

This manuscript has been read and accepted for the
Graduate Faculty in Philosophy in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Sergei Artemov

Date Chair of Examining Committee

John Greenwood

Date Executive Officer

Melvin Fitting

Rohit Parikh

Graham Priest

Samir Chopra

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

TABLEAUX AND HYPERSEQUENTS FOR MODAL AND JUSTIFICATION LOGICS

by

Hidenori Kurokawa

Advisor: Professor Sergei Artemov

In this thesis, we discuss both philosophical and technical issues on proof theory of modal logic and justification logic.

In Chapter 2, we present a view of the foundations of logic, aiming for giving a view of various non-classical logics (called a “structural-reflective view of logic”) and answering the question “are modal operators logical constants?” We present a method of introducing logical constants in Gentzen-style sequent calculi, based on abstract consequence relations. We propose a synthesis of Avron’s, Došen’s, and Sambin et al’s methods of introducing logical constants as positive criteria for a logical constant. In particular, we extend their methods to a certain class of modal logics by adopting the framework of hypersequent calculi and argue that modal operators in these logics are indeed logical constants according to our criterion. In addition, we discuss philosophical repercussions of the method, such as the significance of cut-elimination, the connection of the view with proof-theoretic semantics,

Belnap's criteria of logical constant-hood, and the problem of what a good proof system is.

In Chapter 3, we present hypersequent calculi for some of the strict implication logics and modal logics that are introduced in Chapter 1 and related logics. We show the cut-elimination theorem for these logics and proof-theoretically show correctness and faithfulness of modal embeddings of their superintuitionistic counterparts into these logics.

In Chapter 4, we discuss another application of hypersequent calculi to modal logic. In this chapter, we consider logics that combine Artemov's justification logic and traditional modal logics. We formulate combinations of the logic of proofs LP and traditional model logics S4, GL, and Grz, which are studied from the viewpoint of (either formal or informal) "provability." To handle proof systems for these logics uniformly, we need a proof-theoretic framework that is more general than traditional Gentzen-style calculi. We first introduce prefixed tableau systems and then introduce hypersequent sequent calculi for these logics. We show cut-admissibility for all of these systems via a semantic method.

Acknowledgements

First and foremost, I would like to my express my gratitude to my thesis supervisor Professor Artemov. He not only introduced to me the wonderful subject of justification logics but also gave encouragement and suggestions for the improvement of my thesis. I also owe immensely to Professor Fitting in the intellectual development in my thesis. In some crucial points in my thesis work, I used mathematical techniques developed by Professor Fitting, and he also read technical parts of drafts of my thesis and gave me corrections and suggestions for improvement. I am grateful for Professor Parikh's help for me to develop my background in modal logic. I appreciate Professor Priest's help. He read the philosophical part of my dissertation and gave me both substantial and stylistic improvements. I also thank Professor Chopra for his constant supports for my Ph.D. study in various ways. Although they are not involved in my dissertation project, Professor Hamkins and Professor Kossak helped me develop my basic background in mathematical logic.

I thank members of Computational Logic Seminar. In particular, I am grateful for Walter Dean, Renjune Wang, Thomas Ferguson, Roman Kuznets, Bryan Renne, Evan Goris, Junhua Yu, Jongjin Kim, Evangelica Antonakos, and Karen Kletter for having discussions concerning various topics in logic and philosophy and for their friendship and encouragement.

I am grateful for my parents for their very long-time support for my doctoral study. Last but not least, I also thank Nozomi Kato for her support

for my doctoral study.

Contents

Chapter 3 Strict implications and modal logics 152

Chapter 4 Justification Logics 214

Chapter 1

Introduction

Ever since Gentzen introduced his sequent calculi in the 1930s, Gentzen-style sequent calculi have been widely used in many different areas of proof theory. Some of them have philosophical motivations, others have mathematical motivations, and yet others are motivated by applications in theoretical computer science. In this thesis, we pursue two main themes related to Gentzen's sequent calculi. One of these is to discuss philosophical and methodological foundations of logics based on the framework of sequent calculi and some generalization of sequent calculi. The second is a more technical one, in which we present some proof-theoretic results for the generalization of sequent calculi discussed in the philosophical or methodological part of this thesis.

In order to obtain technical results and philosophical views in the thesis, we use two main tools, namely hypersequent calculi and prefixed tableau systems[1], both of which can be taken to be generalizations of Gentzen's original sequent calculi. Hence, let us briefly explain what kinds of technical tools we use.

In both philosophical and technical parts of the thesis, Gentzen's sequent calculi and their particular (minimal) generalizations called "hypersequent calculi" play an essential role. Hypersequents have a long prehistory before their name was really coined, but so far as the recent literature is concerned, the invention of the calculi as we know them is attributed to Arnon Avron ([10], [14]). A hypersequent is a sequence (or multi-set) of ordinary sequents that are connected by a symbol "|". Hence, a hypersequent looks like this: $\Gamma_1 \Rightarrow \Delta_1| \ldots |\Gamma_n \Rightarrow \Delta_n$. The precise reading of "|" varies depending on what kind of logical systems we are talking about, so we will give a precise interpretation of "|" each time we introduce a new system, although, in general, the intuition behind the notation is disjunctive. Hypersequent calculi are introduced in order to accommodate such non-classical logics as relevant logic RM and some superintuitionistic logics ([10] and [14]). In particular, a theoretical goal that Avron envisaged was to find out elegant generalized sequent calculi that enjoy cut-elimination for these logics. This goal was based on the observation that for some logics, it is hopeless to find a traditional Gentzen-style sequent calculus that enjoys cut-elimination. We had the same situation when we worked on the problem of formulating decent cut-free se-

---

[1] For the philosophical view, the basic proof-theoretic framework with respect to which the view is formulated, is certainly not just a tool, but an integrated part of the view, so the statement here contains some simplification.

quent calculi for logics that we will discuss in the later chapters. This is why we needed to generalize the traditional framework of Gentzen-style sequent calculi. It has turned out that hypersequent calculi work in the cases that we are handling in this thesis.

On the other hand, in our technical investigations (in Chapter 4), prefixed tableau systems play a very important role. The method of semantic tableaux was originated by Beth and developed by Smullyan, but the specific method of prefixed tableaux was introduced by Fitting (the systematic development is given in [57], although the original paper goes back to 1970's.). Prefixed tableau systems are tableau systems in which we have a sequence of positive integers in front of each formula (a prefix). Prefixed tableaux are particularly helpful in formulating some modal logics since it can represent possible worlds by prefixes. Since our foundational view (developed in Chapter 2) is proof-theoretic, we do not have too many foundational issues to be discussed about prefixed tableaux, but we do not mean to degrade the approach by saying that. Our view of prefixed tableau systems is that they are very important *tools*. We use the flexibility of prefixed tableau systems and formulate some relatively complicated systems in justification logics.

Let us move from the method to the contents of the thesis. In pursuing the first of the two main themes stated above, in Chapter 2, we focus on conceptual issues and try to understand various non-classical logics from a uniform perspective of both Gentzen's sequent calculi and hypersequent calculi. Starting from abstract consequence relations, we eventually formulate various cut-free sequent calculi and their generalizations for non-classical logics (substructural logics and modal logics). This constitutes the technical back-up for our view of logic. Along with developing our view, we discuss the issue of what logical constants are and how to introduce commonly used logical constants in a proof-theoretical method. On these bases, we try to answer the main question of the part, namely "are modal operators (decent) logical constants?"

Motivations for asking such a question are almost purely philosophical and the question originally came out of curiosity. It seems a meaningful question, but at least we did not have any clear answer to this question when we asked ourselves the question. Hence, we started taking the question seriously.

Now we try to raise this issue seriously not only out of curiosity but also because there are not too many discussions about this in the literature (model-theoretically or proof-theoretically). There used to be some skepticism against modal logic (Quine's critique against modal logic) or some caveats about its potential susceptibility to paradox (Montague-Kaplan paradox), although the latter is only indirectly related to the issue here since it was intended to give a critique to a so-called "predicate view of modality." However, an even simpler question "are modal operators logical constants?" has been rarely asked. This may be partly because this question hardly directly leads to an idea of a theorem to be proven. Logicians may not primarily be interested in thinking about the nature of logic but try to prove theorems.[2] Another reason may be that the general problem of logical constant-hood (logical constancy) itself is already a difficult problem. There are many discussions on logical constants or logicality in the literature, but as far as we know there are very few philosophical discussions on whether or not modal operators are decent logical constants. Due to the nature of the question, one might think that unless we have a clear idea of logical constant-hood, one cannot answer to the question whether modal operators are logical constants. To such a reaction, we would reply that this is not necessarily the case. An answer to the question may not have to depend on a complete demarcation of logical constants. It suffices to show that certain natural class of constant symbols can be safely said to be "logical constants" (only positive criteria for a "hard core" notion of logical constants are required for this argument) and that modal operators have a feature so similar to them that they deserve being called logical constants. This is exactly our strategy to answer the question.

On the other hand, philosophical debates usually need proponents and opponents of some position (or some dialectical backgrounds such as a controversy). If nobody really opposes to our view, there is no point of pushing our view. Fortunately, there are a few authors who are actually against the view that modal operators are logical constants. Interestingly, the most clearly

---

[2] However, this does not necessarily means that some fundamental philosophical question never leads to a significant technical project. When Frege asked a question "what are numbers?," most mathematicians neglected the question.

stated position about this matter is against the view that modal operators are logical constants ([80]). Thus, we may have some reason to defend our view that modal operators are logical constants.[3]

Our answer to this main question is based on our fundamental view of logic, which is initially motivated independently of the problem of logical constant-hood of modal operators (it was developed due to our interest in the general foundations of logic). We present a view of logic which we named "structural-reflective view" of logic. This is a view that has the following basic features. First, we take Gentzen's sequent calculi (and their minimal generalizations, namely hypersequent calculi) as the fundamental framework to present our own view. Second, we take what is called "Došen's principle" ([45]) seriously. The principle states that various non-classical logics should be considered as structural variants ("structural" in the sense of structural rules in Gentzen's sequent calculi), basic rules governing the behavior of logical constants (called "operational rules") being fixed. We call this a "structural view." Third, we also take the view according to which logical constants are taken as the result of injecting "a part of the the metalanguage into the object language"([148]). We call this "reflective view." There are several preceding works that develop such a view in the literature (includ-

---

[3] In asking a rather "naive" question likes this, it may be helpful to go back before a historical moment when the currently dominant approach to the subject become prevalent and to think about what could have been done if the development had not happened in the way it actually happened. In our case, that is Kripke semantics. Apart from our proof theoretic inclination in studying foundations of logics, in our philosophical part of the thesis, we try to think about modal logics without relying on Kripke semantics as much as possible.

ing Avron [16] and Došen [48]). Among these, Sambin et al [143] is one of the most systematic and comprehensive presentations of the idea in the area of substructural logics. Our view of "reflection" is a synthesis of Došen's, Sambin et al.'s and Avron's views (also, Dana Scott's work added an important twist, as is discussed shortly). Sambin et al. call "the principle of reflection" their fundamental machinery to discuss introduction of logical constants. The name of our view is coined because our view is a combination of these views, i.e., the structural view and the reflective view of logic.

However, our view is a combination of the two views obtained only by slightly modifying both of the views. This slight modification was made in order to extend both of the views to naturally handle modal operators. This is inspired by Dana Scott's relatively underappreciated work [148] done in 1970's. It turns out that Scott's idea of introducing modal operators via strict implications has a close affinity to the view described above. But in order to apply the idea to well-known normal modal logics such as $\mathsf{S4}$ or $\mathsf{S5}$, we need some modifications to both the basic idea of structural-reflective view and Scott's original idea. Our claim here is that according to this structural-reflective view of logic, at least some modal operators in a class of normal modal logics, which can be handled in a proof-theoretically simple way, are "logical constants," provided that some sort of implications (i.e., strict implications) can also be called "logical constants." To justify our thesis, we provide informal arguments to show the reason why strict implications are logical constants (provided other implications are logical constants) and why more common implications (classical, intuitionistic, relevant, and linear implications) are logical constants. Then we argue that accepting these things as logical constants leads to the view that modal operators defined by strict implications are logical constants. This may enable us to (at least partially) answer the fundamental questions "Are modal operators logical constants" and "if so, why are modal operators logical constants?"

Our approach to strict implications and modalities can be taken to be an application of Sambin et al.'s principle of reflection to modal logics[4] that is motivated by Scott's relatively old works on modal logics. Sambin et al. discuss a wide range of logics, but they somehow do not discuss modal logics in their paper. It turns out that extending the method to modal logics requires some twist on the method itself. If one prefers to keep methodological purity, then one may be critical to the idea of extending the method to modal logics. However, we believe that such a direction can be fruitful not only because it can give an answer to the philosophical question, but because that direction may lead to a research program to find out an appropriate framework in which substructural logics and strict implication logics (and hence modal logics) can be uniformly formulated.

In the rest of Chapter 2, we mainly discuss the following issues. First, our structural-reflective view itself does not imply the importance of cut-

---

[4] Avron discusses modal logics in the same paper as he discusses other logical constants [16], but Avron's treatment of modal logics uses a semantic method. By using the semantic method, Avron can talk about local vs. global consequences , etc. which seem to require some semantic method. However, Avron apparently did not consider direct applications of the method to modal logics.

elimination, but our view of cut-elimination plays an important supplementary role in our view of logic. We argue that cut-elimination has not only interesting applications but its intrinsic value. Since it seems that this point is not widely accepted, we give some argument for its conceptual importance.

Second, as a kind of consequences of adopting the structural-reflective view, we will examine more traditional views of proof-theoretically oriented philosophy of logic. We give some comments on proof-theoretic semantics. We discuss two particular issues. One is the issue of coherence between Došen's structural view of logic and so-called "proof theoretic semantics," the latter having played an important role in the philosophical discussions of proof-theoretic foundations of logic. We give a support for the view that the notion of meaning discussed in the literature of proof-theoretic semantics can be divided into two different notions and one of them (called "operational meaning") coheres with the structural view. The second one is an influential view of logical constants that came up as a reply to Prior's "tonk" [131]. Roughly, the view is that if one wants to introduce logical constants by specifying their inferential role via introduction rules and elimination rules, one should do that under the metatheoretic constraints called "conservativeness" and "uniqueness" (whose details we will explain in the relevant sections later). Our overall view about conservativeness and uniqueness is that these notions should play more limited roles in the philosophical discussions on the foundations of logic.

The other topic that we discuss in Chapter 2 is the significance of hypersequent calculi from a broader perspective since our philosophical claim heavily relies on the technical development of hypersequent calculi. We briefly survey the literature of substrucural logics and present some important results about the status of hypersequent calculi in this area. Then we compare hypersequents and other approaches in generalizations of Gentzen's sequent calculi.

In Chapter 3, we present the details of some technical results motivated by the philosophical and methodological discussions in Chapter 2. We do this especially for strict implication logics and modal logics. The main result is cut-elimination for hypersequent calculi of several modal logics and strict implication logics in a uniform manner. Although hypersequents may not dramatically extend the realm of modal logics that can be formulated by using sequent calculi in a cut-free manner, hypersequents are still useful in formulating modal logics that are otherwise difficult to handle via traditional sequent calculi. However, practically, S5 has been almost the only modal logic in the literature that has a hypersequent formulation. In the technical part of this thesis, we have improved the situation to some extent.

We mainly work on modal logics extending S4, S4.2, S4.3, S5. However, since traditional sequent calculi are special cases of hypersequent calculi, we can say that our approach also covers K, T, D, K4, KD4, K45, KD45 in a way, although these have already been known. So, all of them are within the scope of our methodological view. Modal logics that contains B or 5 without containing 4 are difficult to handle from our point of view. But they are naturally so according to our discussions in Chapter 2. Since our official view is to introduce modal operators via strict implications, it would be ideal if we can give such a treatment of modal logics as many cases as possible without introducing any artificiality in their formulations. Thus, we give strict implication versions for modal logics including S4 (for which we formulated hypersequent calculi), i.e., SIS4, SIS4.3, and SIS5. We also give some technical contributions to the studies of hypersequent calculi for logics including S4. For instance, we proof-theoretically prove soundness and faithful of Gödel embedding for these logics. Based on the result, we discuss the relationship among strict implication logics, superintuitionistic logics, and modal logic.

In the literature of proof theory of modal logic, researchers have been interested in how to formulate cut-free proof systems for different modal logics that are already semantically formulated. This is certainly fine from a traditional point of view in the study of modal logic, but there may be another way of taking a look at the subject. For instance, there is no a priori reason why the class of modal logics that can be treated uniformly from a semantical point of view, notably from the point of view of Kripke semantics, and the class of modal logics that can be treated uniformly from a proof-theoretic point of view coincide. Rather than creating some complicated proof systems that can handle all the modal logics that can be naturally treated semantically, we try to give a characterization of modal logics whose proof-theoretic treatment satisfies some reasonable standards as good proof systems. We would like to have an overview of what can be done and what cannot be done by certain proof-theoretic framework and why so. To have such a comprehensive view, covering many logics may be only one of the things that we need to do. As a contribution to such a direction, our method seems to make it possible to reach a reasonable limit of the extension of the class of modal logics that can be naturally obtained proof-theoretically via sequent calculi or hypersequent calculi without using possible world semantics.

In Chapter 4, we apply hypersequent calculi to variants of Artemov's Logic of Proofs (LP) ([6]).[5] LP was introduced in order to talk about explicit proof assertions (written as $t : \varphi$ and read "t is a proof of φ" or "t is an evidence of φ"), rather than provability assertion. A family of logics based on LP is now called "justification logics," and many different logics have been formulated and studied. Notably, combining LP with some other traditional logics (most of which have epistemic, doxastic, or provability interpretation) constitutes an important area of research in justification logics. Several combinations of logics have been proposed and, for some of them, cut-free destructive tableau systems and sequent calculi are formulated and studied. However, for a variety of logics, no tableau systems or sequent cal-

---

[5] Except the fact that our use of hypersequent calculi themselves have some general philosophical motivations, this study is based on more technical interests in both hypersequents and logic of proofs than on philosophical motivations. Hence, we use the conventional term "logic" even though conceptually the status of logical constant-hood of proof terms in the Logic of Proof is not entirely clear at this point. We do need more research in this direction, although we are sure that proof-terms can do more than traditional logical constants or traditional proof-terms in type theory especially from the point of view of "reflection" in our sense. We did not discuss this topic in the thesis since it seemed too premature, but the direction is implicitly suggested in [3].

culi have been proposed yet. The common feature of some of these logics is to have the following axiom schema $\neg t : \varphi \to \Box\neg t : \varphi$ (we call this "mixed negative introspection"). What is interesting about this axiom is that under the interpretation that takes $\Box$ as implicit knowledge and $t : \varphi$ as explicit knowledge with evidence, we can read this formula as a mixed version of negative introspection. This seems to be more plausible than the negative introspection axiom in epistemic logic $\neg\Box\varphi \to \Box\neg\Box\varphi$ (we call it "modal negative introspection"), which is known as the S5 characteristic axiom.[6] However, as in the case of S5, handling negative introspection in a Gentzen-style sequent calculus is not easy and so far there exists in the literature no cut-free sequent calculus that can accommodate this principle. To fill the gap, we formulate prefixed tableau systems and hypersequent calculi for justification logics with mixed negative introspection and semantically prove cut-admissibility for these logics. Since it is difficult to directly construct a countermodel for an unprovable hypersequent, we formulate a prefixed tableau system and construct a translation from prefixed tableau systems to hypersequent calculi. Our goals are to prove cut-admissibility of the hypersequent calculus for combined logics S4LPN (LP + S4 + mixed negative introspection) and S4LP and other combined logics GLA (LP + GL) and GrzA (LP + Grz), where GL stands for Gödel-Löb logic, the fundamental modal logic in the area of provability logic and Grz stands for Grzegorczyk logic

---

[6] We leave the argument discussing why the mixed one is more natural than the modal one to the introduction of the relevant section.

a logic well-known in the area of provability logic as a logic of "provability in PA and true." Note that these projects can have dual values. These can be taken to give proof-theoretically satisfactory proof systems for these justification logics, but these can also be taken as interesting applications of hypersequent calculi.

Chapter 2

A structural-reflective view of logic

2.1 Proof-theoretic foundations of logics

We have seen proliferation of logics these days. These logics are either extension of classical logic by adding items in languages such as modal logic or sublogics of classical logic (or incomparable with classical logic). This seems to (at least potentially) raise a philosophical problem of how we should understand the status of these logics as "logic," i.e. as an instrument of reasoning. In this chapter, we try to give a uniform view of these logics. Our view is outlined as follows.

The fundamental notion of logic is given by abstract consequence relations in the sense of Avron [14]. Differences of various logics should be understood as differences of structural features of them in the sense of Gentzen's sequent calculi (counterparts of the conditions of consequence relations), which is occasionally called "Došen's principle" [45] (we also call this view a "structural view"). Also, logical constants are introduced to the object language as reflections of features of the metalanguage (we call it a "reflective view"). In particular, features of logical constants common to structural variants are determined by the operational rules, which are ultimately derived from "the definitional equations" ([143]). We combine these views with some twists, and the resulting view is called a "structural-reflective view of logic."

While we present the view, we will also discuss what this view can say about the problem of logical constant-hood. We propose the derivability of operational rules from the definitional equations as a positive criterion for being logical constants. Our main claim in this chapter is that, according to the structural-reflective view, some modal operators are decent logical constants. We give a support for this by arguing that some strict implications are decent logical constants. Then, we discuss some consequences of this view in some topics in proof-theoretically oriented philosophy of logic. As one of them, we argue that the notion of "operational meaning" [119] in proof-theoretic semantics is compatible with this structural-reflective view. We also examine the significance of important notions or mathematical results in logic from our point of view. We argue that the issue of characterizing logical constants should be made distinct from that of characterizing "good" proof systems. In the former, the notions of conservativeness and uniqueness should play less important roles than traditionally accepted views have suggested. In addition, we argue that cut-elimination has not only technical but *conceptual* significance in the latter issue.

Traditional sequent calculi have turned out to be too restricted for this view to adopt in order to cover many structural variants of non-classical logics (in particular, to treat different modal logics as structural variants). We need hypersequent calculi to express structural features of different logics. (We only deal with propositional logic as the subject matter here.)

Our view of logic is mostly a combination of already existing ideas in the literature. Thus, this chapter is primarily intended to be an opinionated survey.[1] Also, let us note in passing that here we assume that logic is the science of deduction instead of (logical) truth, as opposed to Fregean view of logic as the science of truth ([54]).

2.2 Abstract consequence relations

Our methodological choice for discussing many different logics naturally requires us to use a flexible framework to formulate many different logics. Abstract consequence relations in Avron [16], characterized by some conditions on multisets of formulas in the object language, are helpful for that purpose. Generalizing Tarski's [163] and Dana Scott's ([151], [150]) work, Avron for-

---

[1] A pluralist view of logic in the sense of Beall and Restall ([22]) is closely related to the theme of this section. However, we do not go into the theme itself. We believe that discussions developed here must be useful to prepare for discussing the theme.

mulates abstract consequence relations. Abstract consequence relations are syntactically formulated, and they are different from model-theoretic (semantic) consequence relations. From this viewpoint, traditional proof systems and also semantic consequence relations are all treated as representations of abstract consequence relations ([16]). In this chapter, we present things as syntactically as possible, although, unlike some radical proof theorists, we are not particularly hostile to semantic methods.[2]

Avron's approach goes back to Tarski's consequence relation (or operator) and Scott's consequence relation (there are variants, e.g. [62]) and generalizes these.

1) A Scott consequence relation for a language $\mathcal{L}$ (a generalization of a Tarski consequence relation to a multiple-conclusion one) is a binary relation between sets of formulas in $\mathcal{L}$ and sets of formulas in $\mathcal{L}$ that satisfies the conditions:[3]

1. Strong reflexivity : if $\Gamma \cap \Delta \neq \emptyset$, then $\Gamma \vdash \Delta$;
2. Monotonicity : if $\Gamma \vdash \Delta$, $\Gamma \subseteq \Gamma'$ and $\Delta \subseteq \Delta'$, then $\Gamma' \vdash \Delta'$;

---

[2] Abstract consequence is less contentious than semantic consequence from a certain point of view. Constructivists such as Prawitz [126] are not satisfied with the model-theoretic notion of logical consequence and suggest a constructivist's notion of logical consequence. Abstract consequence relations can be neutral with respect to model-theoretic vs. constructivist consequence.

[3] In [148], Scott calls this "abstract propositional calculus" since this consequence relation is independent of what kind of object language we have for $\mathcal{L}$. Koslow's structural theory of logic [92] is a development of logic from such an abstract point of view (using "implication relation" as primitive, which is not even a syntactic object). Although we use a similar name in our view, we use more or less the traditional object language(s) of propositional logic.

3. Transitivity (Cut) : if $\Gamma \vdash \Delta, \psi$ and $\Gamma, \psi \vdash \Delta$, then $\Gamma \vdash \Delta$.

2) An Avron consequence relation for a language $\mathcal{L}$ is a binary relation between multi-sets of formulas in $\mathcal{L}$ and multi-sets of formulas in $\mathcal{L}$ satisfying the conditions:

1. Strong reflexivity : if $\Gamma \cap \Delta \neq \emptyset$, then $\Gamma \vdash \Delta$;
2. Transitivity (Cut) : if $\Gamma_1 \vdash \Delta_1, \psi$ and $\Gamma_2, \psi \vdash \Delta_2$, then $\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2$.[4]

Scott generalizes Tarski consequence relation to multi-conclusion one, but Avron further generalizes Scott consequence relation to a one that is neither even monotonic nor sets, only multi-sets.[5] Adding other conditions, we can recover Scott consequence relation (Tarski consequence relation is obtained by restricting the set Δ to the right of turnstile to a singleton of a formula in the language). Avron's consequence relation with monotonicity is called "monotonic," Avron's consequence relation that assumes Γ, Δ are sets of formulas called "regular," and Avron's consequence relation that satisfy these

---

[4] Note that at least in a single-conclusion case, reflexivity and transitivity can be formulated by a single biconditional as follows: $\Gamma \vdash C$ iff $\forall \Delta\ \forall B$, (if $C, \Delta \vdash B$ then $\Gamma, \Delta \vdash B$) ([51]). Then reflexivity and transitivity are converse to each other. Some "duality" holds between these (ibid.).

[5] Our consequence relation does not have to be recursively enumerable or compact, although we consider only recursively enumerable and compact cases. See [47]. The consequence relation invariant under uniform substitution is called "structural" by Avron ([9]). (Here the word has nothing to do with Gentzen's structural rule.) We do not require this as a condition for consequence relation here. See [84].

two conditions called "ordinary." Let us note that we make a distinction between $\vdash$ of an abstract consequence relation and $\Rightarrow$ in a sequent calculus.[6] This view can be called a three-layered view of logic (object language, intermediate (semi-formalized) language ($\Rightarrow$), metalanguage ($\vdash$)). Following Došen [42], we call the intermediate language "deductive-metalanguage." This makes sense since we later introduce a generalized sequent calculi called "hypersequent calculi," where we do not identify $\vdash$ and $\Rightarrow$.

2.3 What is a proof system? Natural deduction vs. sequent calculus

As proof systems that syntactically represent abstract consequence relations, we here consider Hilbert-style axiom systems, natural deduction systems, and Gentzen-style sequent calculi. However, Hilbert-style axiom systems seem to fit better to a view that takes logical truth as the fundamental subject matter of logic, since a Hilbert-style system typically consists of many axioms and few rules. But we do not take logical truths as the main subject matter of logic, so we give substantial discussions only on either natural deductions or (generalized) sequent calculi.[7] The goal of this subsection is to summarize some views in the literature of proof-theoretically oriented philosophy of logic

---

[6] When we refer to a symbol in the prose, we try to avoid using excessively many quotation marks unless we have any particular notational awkwardness.

[7] Also, note that there are some non-classical logics that have no logical truths (e.g., Kleene's strong three-valued logic, first-degree entailment (FDE), [130]), for which Hilbert-style systems are not suitable.

that are based on these systems.

First, let us give a brief overview of these systems ([61], [124], [167], [108]).

(1) Natural deduction systems are the systems in which we have a pair of introduction rule and elimination rule for each logical constant. Ideally, these two kinds of rules constitute a pair of rules for each logical constant (only one logical constant is essentially involved in a step of inference (see, 2.5.)), but sometimes there are discrepancies (e.g., classical negation). Deductions are typically constructed via combination of assumptions and discharging the assumptions (usually taken as sets of formulas), and usually there are no axioms.[8]

(2) (Traditional) sequent calculi use the basic format $\Gamma \Rightarrow \Delta$ ($\Rightarrow$ is an implication in the deductive-metalanguage, although this $\Rightarrow$ and $\vdash$ are often identified) and consist of three items, axioms, structural rules, and operational (logical) rules. The axiom corresponds to reflexivity of $\vdash$. Structural rules correspond to the conditions specified for the consequence relation discussed above for each logic. Operational rules, which are common to different logics, are divided into L rules and R rules. L rules (R rules) are considered as corresponding to elimination (introduction) rules in natural deduction systems, respectively (see 2.5). Also, the outstanding features of sequent calculi are : 1. structural rules are separated from operational rules; 2. (usually) cut is the only rule in which a formula occurring in the uppersequent disappears in the lowersequent. So, by proving cut-elimination, (usually) we can prove

---

[8] Sometimes there are some structural constraints on assumptions in natural deductions.

a formula only by using the subformulas of the formula.[9]

Now we discuss a few views on the issue of which type of systems should be suitable for discussing foundations of logic. Interestingly, we can summarize the views in philosophy of logic by taking a look at how they interpret a passage of Gentzen's.

> "The introduction rules represent, as it were, the 'definitions', of the symbols concerned, and the eliminations are no more, in the final analysis, than the consequences of these definitions." (p.80 [161])

Gentzen uses the word "definition" but with the qualifying phrase "as it were," and this left considerable room for different exegetical possibilities in understanding the passage. The issue is not confined to a purely exegetical one since, based on this passage, substantially different views in philosophy of logic have been developed.

(I) Hacking [80] uses Gentzen's sequent calculus to discuss the issue of what logic is. Hacking literally interprets Gentzen's words as follows: "he had the idea that operational rules actually define the logical constants that they introduce (p.296 [80])," but Hacking himself thinks that this idea needs some qualification. He prefers to say that operational rules do not define logical constants but "characterize" them. However, by considering which conditions would have to be satisfied if they are to be regarded as definitions, Hacking

---

[9] There are some exceptions. Justification logics are among those exceptions (see Chapter 4).

picks out the following criteria that operational rules have to satisfy:[10] 1. conservativeness; 2. eliminability.

Hacking chooses the framework of a sequent calculus to show that in that framework these conditions hold. Hacking focuses on classical logic; hence, he assumes reflexivity, monotonicity and transitivity (and sequents are sets of formulas) for consequence relation ($\vdash$) (hence, this is an ordinary consequence relation in Avron's sense), which Hacking identifies with $\Rightarrow$ in a sequent calculus. Hacking claims, "proving cut-elimination is one ingredient in showing that the operational rules are conservative definitions (ibid.)." (The standard definition of "conservativeness" is given in the item (II).) Also, he shows that the properties of reflexivity and monotonicity for general cases are admissible if we have only those properties for atomic cases. This is what Hacking means by the word "eliminability".

Hacking also argues that requiring eliminability excludes the cases of modal operators from the set of logical constants. Hacking's claim is based on the following observations. In the well-known sequent calculus for $\mathsf{S4}$ [38], weakening using a non-atomic formula is not eliminable (e.g., $\neg\Box(p\wedge\neg\Box p)$[11]). This point is ultimately based on some non-local (context dependent) features of modal operators, i.e. having restrictions on side formulas. For instance, R rule for $\mathsf{S4}$ modality restricts formulas occurring the context to ones whose

---

[10] In modern logical treatment of definition, uniqueness is also an important issue, but Hacking does not extensively discuss the issue of uniqueness. See (II) here and 2.4.8 for uniqueness.

[11] This is so when we use the multiplicative rule for conjunction.

outermost logical symbols are $\Box$ (we call these formulas "modalized"). Hacking's desideratum is that all rules should be local. Hence, modal operators are excluded from the set of logical constants. As a result, since the standard treatment of quantifiers has a similar feature (eigenvariable conditions), Hacking also gives up the standard quantifier rules and uses ω-rule.

(II) Dummett and Prawitz gave the second type of interpretation. On Gentzen's passage, Dummett only says, "he meant that it (an introduction rule in natural deductions) fixes its meaning"([53]), but Prawitz more explicitly explain the idea.

> To develop Gentzen's idea we have thus firstly to state more exactly how the introductions *determine the meaning of the logical constants*; the phrase saying that the introductions represent definitions is clearly not meant to be taken literally. The view that I am taking is that the introductions represent what we may call the canonical ways of inferring a sentence. Other ways of inferring a sentence have to be justified by reducing them to the canonical ways. (p.510 [129])

There is an obvious disagreement between Hacking and Prawitz concerning the exegetic issues on Gentzen's phrase "as it were, definitions." Prawitz's interpretation takes introduction rules as "the canonical ways" to give the meaning of a logical constant and never takes the word "definitions" literally.

Due to the emphasis on the idea that introduction rules give the meaning of logical constants, their approach is sometimes associated with the term "proof-theoretic semantics." The name comes from the fact that they give the meaning of logical constants via a proof-theoretic method (see 2.4.7.). Although their own approaches are significantly different from each other (see 2.4.8.), here we emphasize what is common to Dummett's and Prawitz's idea, since Dummett's and Prawitz's views are close enough to constitute a view that can be well contrasted with the other views. Here we mention the following points, although we do not go into the details of Dummett's revisionism. (1) Dummett introduced the concept of "harmony" influenced by Gentzen's remark and by Prawitz's technical work on normalization of natural deduction for intuitionistic logic; (2) their interpretations are connected with two philosophical issues, i.e., justification of deduction and revising logic (namely, repudiating classical logic and adopting intuitionistic logic).

Dummett introduces the concept of harmony via informal considerations of desirable features of our linguistic practice as follows.

> Given what is conventionally accepted as serving to establish the truth of a given statement, the consequence of accepting it as true cannot be fixed arbitrary; conversely, given what accepting a statement as true is taken to involve, it cannot be arbitrarily determined what is to count as establishing it as true. (p.215 [53])

Dummett tries to clarify the concept by making it precise. Dummett [53] makes a distinction between two different notions of harmony. One is called "intrinsic harmony," and the other is called "total harmony." Let us take any part of a deductive inference (in natural deductions) where, for some logical constant c, a c-introduction rule is followed immediately by a c-elimination rule. Dummett calls it a "local peak for c." When we can eliminate a local peak, then we call c "intrinsically harmonious". On the other hand, "total harmony" simply means conservativeness in the technical sense, which we will now explain together with some background.

Dummett's point on total harmony was introduced in his reply to Prior's arugment in [131] using a strange logical constant "tonk," which was introduced to criticize the claim that any introduction rule and elimination rule determine the meaning of logical constants. Prior's tonk can be formulated in a sequent calculus as follows.[12]

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

By using cut, this immediately leads to the following.

$$\dfrac{\dfrac{A \Rightarrow A}{A \Rightarrow A * B} \qquad \dfrac{B \Rightarrow B}{A * B \Rightarrow B}}{A \Rightarrow B}$$

Since A, B can be any arbitrary formulas, this is absurd. Note that cut is not eliminable in a system that has tonk. This argument is sufficient to refute the above claim, since this shows that arbitrarily chosen rules can determine

---

[12] Prior's direct targets are apparently Popper and Kneale. Also, Prior uses natural deduction, but this is convenient for our later purpose.

the meaning of it in such a way that triviality follows, which is obviously useless.

Dummett apparently accepts Belnap's diagnosis of Prior's tonk in [23], for Dummett identifies the notion of total harmony with that of conservative extension. In [23], Belnap claims that, given an ordinary consequence relation, an introduction of a new logical constant has to satisfy the two conditions: conservativeness and uniqueness (we come back to this issue in 2.4.8.). Here, conservativeness is the standard notion of conservative extension formulated as follows. Let T a theory in an old language L and T' be a theory in an extended language T in the language L'. In this case, T' is called a conservative extension of T if for any $\varphi \in L$, if $T' \vdash \varphi$ then $T \vdash \varphi$. Uniqueness is formulated as follows. Let us call the new constant a (and their counterparts a_1 and a_2) and a formula that contains the new constant φ. Let us call φ_1 and φ_2 respectively formulas in which we replace a by a_1 and a by a_2. Then, the extension of T by a (called T_a) is unique if φ_1 and φ_2 are interderivable in the theory of the language extended by both a_1 and a_2 (called $T_{a_1 a_2}$).

However, there seems to be a tension between the technical notion of conservativeness and an informal concept of harmony that Dummett takes to be important based on epistemological considerations (see 2.4.8.). Here we mention in passing that the tension is produced by Dummett's adoption of the traditional notion of conservativeness, which may not fit well with

Dummett's revisionist program of logic.[13]

On the other hand, the notion of intrinsic harmony is essential in his proof-theoretic justification of deduction in [53]. Dummett takes introduction rules in natural deduction systems in intuitionistic logic as "self-justifying." It is "self-justified" in the sense that it does not require any such justification as mentioned in the quote from Prawitz (does not require reductions to the canonical ways). Roughly speaking, proof-theoretic justification of deduction in Dummett's sense is like a reduction procedure used in Prawitz's normalization theorem for natural deduction of intuitionistic logic[14] ([53]), which is essentially the operation of leveling local peaks (Prawitz called this "inversion principle"). This is where the notion of intrinsic harmony appears in Dummett's justification of deduction. Normalization implies that logical constants in intuitionistic logic are intrinsically harmonious. But classical negation is not so. Normalization for classical logic uses more complicated ideas than leveling local peak. Lack of harmony in classical logic is one of the reasons why they take the possibility of revising logic seriously.

(III) Michael Kremer [94] takes a middle way between Hacking and Dummett. Kremer holds the view that rules in Gentzen's sequent formulations of logic give the meaning of logical constants or define them (in a certain lim-

---

[13] Dummett himself makes qualifying remarks about conservativeness, such as "this notion is in a high degree relative to the context, that is, the base theory to which the addition is being made." ([53]. See also [52].) However, Dummett does not go into discussions on the qualification.

[14] A normal (and canonical) proof is, roughly speaking, a proof without detours (or "peaks") consisting only of applications of introduction rules.

ited sense). Kremer suggests modifications to both Dummett's and Hacking's position as follows.

> The basic idea is to take the left and right rules in a sequent calculus as 'defining' or 'giving the meaning of' the introduced symbols. ... The view under consideration, however, takes the left and right introduction rules jointly to give the meaning of the introduced symbols without positing any order of priority between the two. (p.54, [94])

In spite of their superficial similarities, Kremer's approach and Dummett-Prawitz approach radically differ, since Kremer is not a revisionist and he separates the issue of classical vs. intuitionistic logic from that of the meaning of logical constants. Still, Kremer takes the notion of conservativeness as playing an important role in introduction of logical constants (but by appealing to a multiple-conclusion sequent calculus for classical logic and its cut-elimination (see 2.4.7.)).

Here is a summary of the discussions above.

	system	logic	key notion(s)
Hacking	sequent calculus	classical	conserv. + elim.
Dummett	natural deduction	intuitionistic	total/intrinsic harm.
Prawitz	natural deduction	intuitionistic	inversion principle
Kremer	sequent calculus	classical	conservativeness

We have general comments on these interpretations of Gentzen's remark. Overall, we are not quite satisfied with their works because they only consider classical logic and intuitionistic logic. Dummett-Prawitz program is interesting as a study of foundations of intuitionistic logic, but our primary goal here is to understand many different logics uniformly, possibly extending Gentzen's framework for a logical system (note that Gentzen himself treated only classical and intuitionistic logic). Although fully examining the adequacy of natural deduction as a fundamental framework for a logical system is beyond the scope of this work, we can at least say that natural deduction does not fit our purpose here, since it works particularly well only for intuitionistic logic.[15]

Like Hacking and Kremer, we prefer sequent calculi as a basis for discussing logics, but for a reason different from theirs. They deal with only classical logic, for which a multiple-conclusion sequent calculus is suitable, but they do not pay much attention to structural rules. However, we need to do that, since we aim to have a uniform view of different logics, most of which are differentiated by structural rules.

More importantly, we disagree to Hacking's exclusion of modal operators

---

[15] See Garson's work [65] that uses valuational semantics to characterize natural deduction for intutionistic logic. Neil Tennant defends his version of intuitionistic relevant logic, which is different from traditional relevant logics, along the line of Dummett-Prawitz's program ([165], [166]).

from a class of logical constants. We will argue that there is a sense in which at least some modal operators are logical constants. Also, concerning cut, we are in agreement with Hacking in thinking that cut should be eliminable, but we do not take conservativeness to be our main reason for thinking so.

Although their works are important since they take a proof-theoretic point of view and focus on some meta-theoretic notions such as conservativeness in order to discuss philosophical issues (not only for technical purposes), the notions used in their discussions are highly context-sensitive, and we have to be careful in handling these concepts. This point is relevant here since we consider many non-classical logics. In the next section, we introduce a philosophical view of logics that can handle different non-classical logics. We will revisit the issues of the meaning of logical constants, conservativeness, uniqueness, and harmony, from such a broader perspective.

2.4 From Došen's principle to a structural-reflective view of logic

The approaches that are discussed in the last section do not take into consideration the existence of many non-classical logics. In contrast to these approaches, Kosta Došen takes into account many different non-classical logics by focusing on structural rules in sequent calculi. The idea can be summarized by the phrase: **logical constants as punctuation marks** [45]. Here we discuss this view and our own modifications of it. Before going into substantial discussions of the view, let us mention our plan in this section. After presenting Došen's, we introduce our own view ("the structural-reflective view"). We extract a positive criteria for logical constant-hood from one of the main constructions in the formulation of the view.[16] We argue that this can be naturally extended to cover modal logics (which contains some extension of a proof-theoretic framework). Hence, we claim that significantly many modal operators that can be handled by our method are logical constants, giving some philosophical argument to defend the claim. Then, we discuss supplementary issues such as the significance of cut-elimination in 2.4.6. (which plays a crucial role in our dialectical position) or consequences of the entire "structural-reflective view" (2.4.7. and 2.4.8.)

Let us start discussing Došen's view. Došen's view assumes the following two fundamental principles:

> [A] logic is the science of deductions;
>
> [B] basic formal deductions are structural deductions.

Under these assumptions, Kosta Došen proposes the two theses in [42].

> [I] A constant is logical if, and only if, it can be ultimately analyzed in structural terms.
>
> [II] Two logical systems are alternative if, and only if, they differ only in their assumptions on structural deductions.

---

[16] The issue of logical constant-hood arises only two places in the discussions in 2.3., i.e. Hacking on modal operators and Dummett on tonk. This is probably because they argue for only intuitionistic or classical logic. From our point of view, this issue has more substantial significance since we try to handle many different logics.

The latter is commonly identified as "Došen's principle," but both principles require detailed explanations. $[I]$ contains a phrase "ultimately analyzed in structural terms." The word "structural" simply comes from the structural rules in Gentzen's sequent calculus ($[II]$ should also be understood from this terminology), but the word "analyzed" needs to be explained. According to Došen, an expression is ultimately analyzed if and only if it is either analyzed or it can be analyzed in terms of analyzed expressions ([45]). This definition depends on what the word "analyze" means. Thus, let us give an exposition of Došen's account of analysis. Suppose that we have an expression α of a language L to be analyzed. Then we specify a language M that does not contain α but in which we want to formulate the analytic equivalent of α (analysans). An analysis should satisfy the following conditions.

> (1) An analysis consists in establishing that a sentence A in M plus α, in which α occurs only once, is equivalent to a sentence B in M (p.369, [45])
>
> (2) From the equivalence of (1), and from the understanding of M and L minus α, we can infer every sentence of L which is analytically true in L and no sentence of L which is not analytically true in L. (ibid.)
>
> (3) The expression α_1 and α_2 can receive the same analysis if and only if α_1 and α_2 have the same meaning. (ibid.)
>
> (4) The language M should be more basic than the language L.

(ibid.)

Let us first note that in Došen's terminology "analytically true in L" means "true in virtue of the meaning of the expressions of L"([45]). (1) is about the form of analysis, (2) states that an analysis must be sound and complete in some appropriate sense (a full explanation requires what we mean by "meaning"), and (3) states that an analysis characterizes uniquely (up to synonymy) the expression analyzed. In (4), Došen states the priority of the metalanguage. Since these conditions for Došen's notion of analysis are abstract, to further explain what Došen means by "analysis," we will cite an illustration. But this itself requires some preparations.

2.4.1 Structural rules, operational rules, and logical constants

We present Došen's structural analysis of logical constants and then our own variant of Došen's approach. We first present structural rules and Došen's versions of "operational rules."

2.4.1.1 Structural rules

One of the main points of Došen's theses is to separate the roles that the structural rules play and those that operational rules play in various logical systems. Standard structural rules that go back to Gentzen [61] are as follows.[17]

$$\text{Weakening} \qquad \textbf{LW}\ \frac{\Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta} \qquad \textbf{RW}\ \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, A}$$

$$\text{Contraction} \qquad \textbf{LC}\ \frac{A, A, \Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta} \qquad \textbf{RC}\ \frac{\Gamma \Rightarrow \Delta, A, A}{\Gamma \Rightarrow \Delta, A}$$

$$\text{Cut} \qquad \frac{\Gamma_1 \Rightarrow \Delta_1, A \qquad A, \Gamma_2 \Rightarrow \Delta_2}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2}$$

Here let A, B, C … be schemata for formulas of an object language, and let Γ, Δ, … be schemata for finite multisets of formulas of an object language. Let us briefly discuss what schemata mean. Došen has a general distinction between constants and schematic expressions for any language [42]. We give up a systematic presentation of the details of Došen's discussions on the notions of schemata and constants, since they are too tedious but apparently a precise reconstruction of the commonly accepted notions. Practically, all we need here are (1) schemata in the deductive metalanguage to talk about certain patterns of formulas in the object language, for which we have certain rules of (uniform) substitution w.r.t. the formulas of the object language, and (2) constants in the object language, which, roughly speaking, do not allow substitutions. Some of these constants deserve being called "logical constants."

The structural rules presented above are given only as paradigmatic ex-

---

[17] If we start with sequence of formulas, we use structural rule "exchange."

amples. It is difficult to give an exhaustive list, since there is no uniform agreement about what structural rules are (see 2.5). However, Došen's definition may be a good starting point to present our view. Došen defines structural systems in [44] as follows.

Definition 2.4.1 *An expression of a language $\mathcal{L}$ is essential for a system S iff it occurs at least once in a rule, axiom or axiom schema by which S is presented, or it occurs, or is referred to, at least once in a proviso of a rule, axiom or axiom-schema by which S is presented. A system in the deductive-metalanguage is structural iff no constant of the object language is essential for it.*

More briefly, Došen states, "we shall also say that rules, deductions, etc., are structural, whenever they do not involve any constant of the object language" ([44]). It is crucial to note that the condition that no constant of the object language is involved is essential for Došen's position.

2.4.1.2 Operational rules and logical constants

We present Došen's "operational rules." Strictly speaking, Došen does not use operational rules but double-line rules [45]. (For traditional operational rules, see 2.5.1). Došen presents his systems as packages of structural rules and double-line rules.

Double-line Rule for $\to$: $(\to)$ $$\frac{\Gamma, A \Rightarrow \Delta, B}{\overline{\Gamma \Rightarrow \Delta, A \to B}}$$

$$\text{Double-line Rule for } \wedge : (\wedge) \quad \frac{\Gamma \Rightarrow \Delta, A \qquad \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \wedge B}$$

$$\text{Double-line Rule for } \vee : (\vee) \quad \frac{\Gamma, A \Rightarrow \Delta \qquad \Gamma, B \Rightarrow \Delta}{\Gamma, A \vee B \Rightarrow \Delta}$$

The double-line notation means that we can infer both downwards and upwards. In the following we use notations $\downarrow$ ($\uparrow$) to mean the downward (upward) direction of the rules. The most important feature of Došen's double-line rules is that they are fixed throughout different logics that Došen formulates by using these rules. In this sense, classical logic, intuitionistic logic, and a fragment of relevant logic R can be variants in terms of structural rules. The rule for $\rightarrow$ is classical, but Došen shows that by restricting the weakening rule on the right to a case the lowersequent has one formula on the succedent we get intuitionistic logic (dropping weakening on the succedent will give minimal logic). Dropping weakening on the both sides, we can formulate a fragment of relevant logic R (without distribution), based on the common rules governing logical constants ([45]).[18] Hence, this is a good illustration of Došen's principle.[19]

This way of differentiating intuitionistic logic and classical logic is not the traditional one. These are distinguished by using the number of formulas allowed to occur on the succedent. As is well-known, if we allow at most

---

[18] Note that this classical rule is intuitionistically unsound (w.r.t. Kripke semantics). Nonetheless, using cut-free formulation, we can show that we do not prove any formula that is intuitionistically invalid. In this formulation, the mentioned relevant logic must be a subsystem of intuitionistic logic.

[19] In his technical work [48], Došen gives a more comprehensive presentation of double-line rules.

one formula to occur on the succedent (using the standard classical rules otherwise), then we obtain intuitionistic logic. Like this case, we can differentiate logics via contexts in many cases. Although we mostly follow Došen's structural view, we include the notion of contexts into the class of structural features of a proof system, not only structural rules. Hence, the notion of "structural variants" can be defined as follows: sequent calculi (or hypersequent calculi) are structural variants to one another if they share operational rules but have different structural rules or contexts.

Došen does not discuss the relation between double-line rules and traditional operational rules in [45], although he does so in a technical paper [48]. This may indicate that Došen may not want to let his philosophical points rest on the relation between these. However, from our point of view, the link between double-line rules characterizing logical constants and traditional operational rules in sequent calculi are extremely important, since the double-line form is not sufficient to formulate a cut-free sequent calculus. Downward rules are the same as traditional operational rules, but upward rules do not have the ordinary forms of operational rules. We later show how to derive traditional operational rules from double-line rules.

2.4.2 Došen's structural analysis examined

With these backgrounds, we can understand Došen's illustration of a structural analysis of logical constants. Došen uses the case of implication " $\rightarrow$ ".

> In the analysis of implication given by $(\rightarrow)$ the language L is the language of propositional, or first-order, logic, and M is the deductive metalanguage in which we speak about structural deductions; more precisely, M is the language of structural sequents. The sentence A is the lower sequent of $(\rightarrow)$, B is the upper sequent of $(\rightarrow)$. Since the double line stands for an equivalence our analysis has the form prescribed in condition (1). In [44] it was shown that $(\rightarrow)$ serves to characterize implication, soundly, completely and uniquely, in classical, intuitionistic, and relevant logical systems; i.e. conditions (2) and (3) are satisfied. Finally, we suppose that our structural analysis of implication satisfies condition (4), since we suppose that the language M of structural sequents is more basic than the language L of propositional, or first-order, logic. (p.373, [45])

Let us now start examining Došen's analysis. The fundamental format of Došen's analysis, namely double-line rules plus structural deductions, is agreeable. However, we propose some modifications of a few items in Došen's structural analysis. Concerning (2) and (4), we do not have any objection, although (2) is vague without any explanation of "meaning." Depending on what we mean by "meaning" there may be different ways of considering correctness and adequacy. Since Došen does not exclude a possibility of having model-theoretic semantics for the logics discussed, this can mean semantic soundness and completeness. But it could be more than this. Since the condition is abstract, there must be some specific way of interpreting the word "meaning" that one can agree to (even when there is some disagreement about particular details). Also, note that our agreement to (4) has consequences to our view of meaning. We come back to those issues on meaning in 2.4.7.

(3) is essentially a uniqueness condition in the sense discussed in 2.3. However, we do not think that uniqueness of a logical constant, as formulated in 2.3., has to be necessary. In [44], Došen argues that this condition is highly sensitive to formulations (proof systems) of a logic. For instance, Došen showed that Hilbert-style systems cannot uniquely characterize intuitionistic implication, but a sequent calculus for intuitionistic logic does, and a sequent calculus for S4 cannot uniquely characterize S4 modal operator, but Došen's higher-level sequent calculus does [44] (see the appendix for the system). Uniqueness of a logical constant may be a desirable property for some proof system, but it is unclear why this condition has to be uniformly required for every logical constant. Since the issue involves philosophical discussions, we will further discuss it in 2.4.8.

The condition (1) states that α can occur only once. Došen suggests a possibility of relaxing this condition, namely analyzing a sequence of α, but Došen probably does not mean that α occurs also on the other side of the double-line rule. Indeed, violation of condition (1) in this way seems to be problematic for two reasons. One is that this spoils purely schematic nature of logical rules. The second is that an analysis whose analysans contains an analysandum may look circular. However, we later argue that despite these worries, we should loosen the conditions (at least in some cases). In particular, concerning the second point, we argue that this seeming circularity is not harmful and that a logic with rules formulated in this way can even have some conceptual advantage compared with intuitionistic logic (see 2.4.3.2). Accordingly, we later loosen Došen's notion of structural deductions. (This loosening may look too far away from Došen's view to keep using the term "analysis". We do not mind giving up the term "analysis" and use "characterization." See 2.5.2.)

To avoid possible misunderstanding, let us now discuss what Došen's structural analysis is *not*. This will be helpful since it is not an easy task to understand exactly what Došen's structural analysis is. (Došen [45] compares his analysis with the analysis of the notion of computation in Church's thesis.) Briefly put, Došen argues that his structural analysis gives neither the definition nor the meaning of a logical constant. Concerning the meaning, we need to discuss it against the background of proof-theoretic semantics, so we put off the discussion until 2.4.7. and concentrate on the issue of definition.

Došen argues that his analysis does not give a definition since his structural analysis does not satisfy the two additional conditions that a definition has to satisfy.

> The definitional equivalence should enable us to find for every sen-

> tence of M plus α a sentence of M with the same meaning.(p.369, [45])
>
> Every sentence of L minus α, which is analytically true in L, is analytically true in L minus α. (p.370, [45])

The first one of the two conditions is called "Pascal's condition." This amounts to the requirement of eliminability of the defined expression by its definiens. The second condition for definitions is "the conservativeness condition." Došen remarks that for ($\rightarrow$), Pascal's condition and conservativeness do not necessarily hold for every structural analysis,[20] although conservativeness holds for classical, intuitionistic, and relevant logic. Došen gives a cautious comment on conservativeness.

> Our replacement of the term "definition" by the term "analysis" in the above program is not merely a verbal move. One substantial difference is that the requirement of conservativeness ceases to play the role it has to play in the former program. (ibid.)

---

[20] Došen [45] gives the following examples. The former fails because in a single-conclusion sequent-system we may be unable to eliminate implication from the sequent of the form $A \rightarrow B \vdash C$ (consider intuitionistic implication, which is not definable from disjunction and negation). Conservativeness fails in some case. If we add the double-line rule ($\rightarrow$) to a sequent-system with weakening on the right only, we can derive weakening on the left, and this makes L nonconservative with respect to L minus $\rightarrow$. For instance, $A \Rightarrow A$ implies $A \Rightarrow A, B \rightarrow A$. Then, $A, B \Rightarrow A, A$. Thus, $A, B \Rightarrow A$, which is like weakening.

We are sympathetic to Došen's views, which let both eliminability and conservativeness go. If an analysis is to be a definition, it has to satisfy conservativeness, uniqueness, and eliminability. However, these conditions may be too strong. Hence, it is correct to avoid saying that Došen's analysis gives a definition. In 2.4.8., we will later give some examples of non-conservative cases that satisfy the conditions of Došen's analysis.

Before moving on, let us give a few more general comments on the scope of Došen's view. First, Došen's principle as originally formulated has a limitation, since many systems that are called "logics" currently have no sequent calculi. This problem is partially solved by generalizing sequent calculi (see 2.5 and Chapter 3).

Secondly, Došen does not explicitly address a more traditional question of demarcation of logic, but it is natural to ask what we can say about this from Došen's point of view. Concerning this issue, Došen comments as follows.

> In general, thesis [I] can effectively be applied to show that a constant is logical: if it is ultimately analyzed in structural terms we can claim that it is logical. On the other hand, this thesis does not seem to be effective in showing that a constant is not logical, for it is not clear on what grounds we could claim that an ultimate analysis in structural terms is impossible. (p.376, [45])

This means that his conditions give only a positive test. Although our own proposal contains some modifications of Došen's analysis, we are also satisfied with a positive test, since we may have only a vague idea about the boundary of logic.[21] Moreover, the positive test is sufficient for our two purposes here: presenting a coherent view of logical constants covering reasonably many cases; arguing for the thesis that modal operators of most (normal) modal logics are logical constants.

2.4.3 The principle of reflection for logical constants

In this section, we present a view of logical constant that will be combined with Došen's structural view in section 2.4.4. We try to synthesize the views of the three authors, Sambin et al. [143], Došen [45], [48], and Avron [16]. The basic points are 1. we separate the structural part and the operational part, 2. we start from bi-conditional characterizations of logical constants that are invariant throughout different logics, 3. we derive the traditional operational rules in (cut-free) sequent calculi,[22] 4. we do so by using a fixed meta-theoretical framework. Their views can be summarized in the table

---

[21] For instance, [71] argue that the problem is unsolvable (but not in a mathematical sense). In the tradition of model theory, which goes back to [164], whether or not one expression is logical is not considered to be determined by the notion of consequence itself, but once it is determined, semantic consequence relations determine validity. It seems too ambitious to try to give a necessary and sufficient condition for being logical constants.

[22] Došen reports the derivability only in [48]

below.[23] (Technical terms here will be explained later.)

	biconditional	deriving operat. rules	metatheory
Sambin et al.	defin. equation	solving the equation	metalinguistic link
Avron	Yes, but no name	Yes, but no name	abstract conseq.
Došen	double-line rule	Yes, but no name	deductive-metalang.

Although Avron and Došen do almost the same thing (except implication), Sambin et al. have the most detailed discussions on the derivations of operational rules. So, we first follow their formulations and terminologies and point out how we should modify their setting according to our own theoretical interest.

Sambin et al. formulate a logic called basic logic (BL) that is given as a minimal system which can be extended to linear (relevant), BCK, intuitionistic, quantum, and classical logic by relaxing contexts and adding structural rules. BL has no structural rules, so BL looks like multiplicative and additive linear logic (MALL) except that BL has no context on the side in which operational rules are applied. This feature is called "visibility," which is one of their theoretical slogans, i.e., symmetry, visibility, and reflection. We are not interested in visibility (or symmetry) here, so we do not have to restrict our sequents in this regard.[24] We rather concentrate on the issue of reflection.

---

[23] Their theoretical goals are different. Došen aims for a philosophical analysis of logical constants, Avron has a practical interest in having a uniform framework that works for many logics, and Sambin et al. tries to have minimal logic common to different logics (including quantum logic).

[24] Except their use of co-implication, roughly, their notion of symmetry means just bal-

(For the basic system B of BL, see the appendix.)

Sambin et al. explain what is called "the principle of reflection" as follows.

> The common explanation of the truth of a compound like $A\&B$ is that $A\&B$ is true if and only if A is true *and* B is true. In our terms, a connective $\circ$ between propositions, like & above, reflects at the level of object language a *link* between assertions in the meta-language, like *and* above. The semantical equivalence
>
> $A \circ B$ is true if and only if A is true link B is true,
>
> which we call *definitional equation* for $\circ$, gives all we need to know about it. $A \circ B$ is semantically *defined* as that proposition which, when asserted true, behaves exactly as the compound assertion *A true link B true*. The inference rules for $\circ$ are easily obtained by *solving* the definitional equation, and they provide an explicit definition. We then say that $\circ$ is introduced according to the *principle of reflection*. (p.980, [143])[25]

---

anced occurrences of formulas on the antecedent and the succedent. We do not put emphasis on it, either.

[25] Sambin et al. do not explicitly discuss the conditions for "explicit definition" in the strict sense discussed in, say, Došen [45]. However, our modal cases do not satisfy some minimal conditions for explicit definitions (e.g., occurring only one side of a defining biconditional). We rather loosen the conditions for logical constants and give up this requirement (see 2.4.3).

Here "solving the equation" essentially means deriving operational rules in the ordinary operational rules in sequent calculi. To do that without going into "circularity"[26], they put priority in the metalanguage by stating, "it is the meta-level which comes first, and based only on it the formal system is built up."(ibid.) (Note that this is the same as Došen's condition (4).) They make explicit the meta-theoretical framework as follows.

1) Let A, B, C ... denote propositions.[27] A proposition A must be kept distinct from the assertion about it, and an assertion is written as "*A is*" to keep this neutral among A is true, available, measured, etc.

2) Complex statements built up from atomic assertions via some metalinguistic *link* are introduced. They are called "compound assertions." The compound assertions used in any sequent calculus can be seen as obtained from atomic assertions by means of only two metalinguistic links, namely *and* and *yields.*

A conjunction of atomic assertions C_1 *is and*... *and* C_n *is* is abbreviated as $C_1, \ldots, C_n$, where comma takes the place both of *and* and of *is.* In $\Gamma \vdash \Delta$, we write Γ for any conjunction of atomic assertions, either empty or $C_1, \ldots, C_n$. Similarly, for Δ and $D_1, \ldots, D_m$. Then the meaning of a sequent $\Gamma \vdash \Delta$ is (C_1 *is and* ... *and* C_n *is*) yields (D_1 *is* and ... *and* D_m *is*). The $\vdash$ is a shorthand for *yields.* The meaning of one premise inference

---

[26] This circularity is different from the one discussed in relation to Došen's condition (1).

[27] Sambin et al. do not discuss what propositions are.

$\dfrac{\Gamma \vdash \Delta}{\Gamma' \vdash \Delta'}$ is (Γ *yields* Δ) *yields* (Γ' *yields* Δ'). Namely, Gentzen's horizontal bar is also a shorthand of *yields.* Also, a blank space between two premises of a two-premise rule is also taken as *and* link.

And in the outermost link is the same as conjunction. When *and* is inside the scope of *yields* (say, (A *is and* B *is*) $\vdash \Delta$), nothing is assumed on this *and* except that it behaves well with respect to composition of derivations, as explained below. The same comment applies to *and* inside the scope of turnstile, at the right, and no link interpreted as disjunction is needed. When *yields* is the principal link, $A \vdash \Delta$ means that from A *is* we can get to know that Δ. Sambin et al. formulate the following sequent rules (based on the meaning they assigned to $\vdash$).

The identity rule: $A \vdash A$

Composition on the left $\dfrac{\Gamma \vdash A \qquad \Gamma' \vdash \Delta}{\Gamma'(\Gamma/A) \vdash \Delta}$ on the right $\dfrac{\Gamma \vdash \Delta' \qquad A \vdash \Delta}{\Gamma \vdash \Delta'(\Delta/A)}$

Here $\Gamma'(\Gamma/A)$ means the replacement of one occurrence of A by Γ in Γ'. Commonly used cut (transitivity) does not hold in BL, and cut is broken into the two rules of composition. (We do not adopt these, but we present these for reference.) Note that Sambin et al. do not make any distinction between $\vdash$ and $\Rightarrow$. We keep using consequence relation $\vdash$ (multiple-conclusion or single-conclusion) until we officially introduce our (hyper)sequents. Also, note that Sambin et al.'s account of $\vdash$ cited above is epistemic (p.983 [143]), but our official account (using abstract consequence relations) of $\vdash$ is not.

We will give some illustration of derivations of operational rules from definitional equations, but we do not follow exactly the same meta-theoretic setting as Sambin et al's. Unless we say otherwise, we assume that the underlying notion of "and" is that formulas are connected as a multi-set, and that "yield" is based on the same (deductive) notion as $\vdash$ in Avron's consequence. We do this because their *yield* is not transitive but we assume transitivity. Also, Sambin's approach does not justify exchange rule, so they later add it. But it should be built in the meta-theoretic framework from the outset.

Due to the issue raised only in the derivation of implication, we treat implication and other cases differently. Partly following Sambin et al's terminology (which is originally not used for classifying logical constants), we call the two cases as follows: 1. logical constants that do not reflect $\vdash$; 2. logical constants that reflect $\vdash$.

Logical constants that do not reflect $\vdash$ are disjunction and conjunction. We have both primitive and derivative cases for logical constants that reflect $\vdash$. Implication is treated as primitive, and a modal operator and a negation are derivative.

2.4.3.1 Logical constants that do not reflect "$\vdash$"

Here we show the promised derivations for some representative cases. The definitional equations for multiplicative conjunction and disjunction are given as follows. The definitional equation gives the usage of "," and $\otimes$ on the antecedent and "," and $\oplus$ on the succedent.[28] Note also unlike Sambin et al.'s original derivation, we use Γ and Δ on the same side of $\vdash$ in introducing the new symbol.

Multiplicative (Internal) conjunction : $\Gamma, A, B \vdash \Delta$ iff $\Gamma, A \otimes B \vdash \Delta$

Multiplicative (Internal) disjunction : $\Gamma \vdash \Delta, A, B$ iff $\Gamma \vdash \Delta, A \oplus B$

Definitional equations for additive conjunction and disjunction are as follows.

Additive (Combining) conjunction : $\Gamma \vdash \Delta, A \wedge B$ iff $\Gamma \vdash \Delta, A$ and $\Gamma \vdash \Delta, B$.

Additive (Combining) disjunction : $\Gamma, A \vee B \vdash \Delta$ iff $\Gamma, A \vdash \Delta$ and $\Gamma, B \vdash \Delta$.[29]

Combining this with different combinations of structural rules (contexts) in sequent calculi, we can formulate different logics in a uniform framework.

We illustrate derivations of operational rules, using $\otimes$. By the definitional

---

[28] Avron calls them "internal" and "external." Relevantists call them "intensional" and "extensional." We follow Paoli's notations for conjunction and disjunction [118].

[29] Sambin et al. are explicit in stating that "if", $\vdash$ and horizontal line are all the same (namely, *yields*) but the others are not. Hence, the exact difference between Sambin et al.'s and Avron's 'iff' and Došen's is not entirely clear, but no important issue seems to rest on this point.

equation $\Gamma, A, B \vdash \Delta$ iff $\Gamma, A \otimes B \vdash \Delta$, we can immediately derive the following. The former of the rules is called "formation rule"; the latter "implicit reflection."

$$\otimes\text{-formation } \frac{\Gamma, A, B \vdash \Delta}{\Gamma, A \otimes B \vdash \Delta} \qquad\qquad \text{implicit } \otimes\text{-reflection } \frac{\Gamma, A \otimes B \vdash \Delta}{\Gamma, A, B \vdash \Delta}$$

The first one is already the desired rule, but the second is not. To obtain the desired rule in a cut-free system, we start from $\Gamma \vdash A, \Delta$ and $\Gamma' \vdash B, \Delta$. By using this, Sambin et al. first derives the following principle equivalent to implicit $\otimes$-reflection.

$$\text{axiom of } \otimes\text{-reflection} \qquad A, B \vdash A \otimes B$$

This is obtained by using identity rule for $\vdash$, i.e., $A \otimes B \vdash A \otimes B$. In Sambin et al. this is one of the two rules at the meta-level. (For Avron, this is an instance of strong reflexivity.) Using cut, we can derive the desired rule.[30]

$$\text{explicit } \otimes\text{-reflection} \qquad \frac{\Gamma \vdash A, \Delta \qquad \Gamma' \vdash B, \Delta'}{\Gamma, \Gamma' \vdash A \otimes B, \Delta, \Delta'}$$

This derivation can be given by the following two applications of cut.

---

[30] Note that the obtained formula is not invertible.

$$\dfrac{\dfrac{\Gamma \vdash A, \Delta \qquad A, B \vdash A \otimes B}{\Gamma, B \vdash A \otimes B, \Delta} \qquad \Gamma' \vdash B, \Delta'}{\Gamma, \Gamma' \vdash A \otimes B, \Delta, \Delta'}$$

The $\otimes$-formation and explicit $\otimes$-reflection are the operational rules for $\otimes$ in multiplicative and additive linear logic (MALL) presented in 2.5. These derivations are taken as solving the definitional equation, since the connective $\otimes$ is characterized by the $\otimes$-formation and the explicit $\otimes$-reflection. Here they say "$\otimes$ reflects *and* at the left of the turnstile"([143]). The case of multiplicative disjunction $\oplus$ (the definitional equation is $\Gamma \vdash A, B$ iff $\Gamma \vdash A \oplus B$) is similar, and they say "$\oplus$ reflects *and* at the right of the turnstile." (Here, not only "," but *and* are interpreted differently in the antecedent and the consequent.)

For additive (context-sharing) cases, things go similarly except that *and* does not occur in the scope of *yields*, $\vdash$. This is obvious from the definitional equation for additive conjunction (disjunction) $\wedge$ ($\vee$). They say "$\wedge$ reflects *and* at the right of the turnstile", and "$\vee$ reflects *and* at the left of the turnstile"(ibid.). The metalinguistic tools we use in the case of $\oplus$ and $\vee$ are still "," and *and*. The difference between multiplicative and additive connectives is that $\wedge$ and $\vee$ reflect a link *and* which is not in the scope of the *yields* and $\otimes$ and $\oplus$ reflect a link *and* which is in the scope of *yields*.

2.4.3.2 Logical constants that reflect "$\vdash$"

a. Implication (conditional):

Unlike the cases of $\otimes$, $\oplus$, $\wedge$, and $\vee$, there is a subtle issue about introducing implication via the principle of reflection. Although Sambin et al. introduced the term "the principle of reflection" as solving the definitional equation, we here use the word "reflection" a little loosely due to this problem. We identify the fundamental idea as "injecting a part of the metalanguage into the object language" (Scott[148]). This itself may not guarantee the derivability of operational rules, but we are interested only in cases where they are derivable. We discuss Sambin et al.'s problem after presenting our derivation of rules for implication.

We first present Avron's derivations of operational rules from what we call definitional equations.[31] (The following two cases are structural variants via contexts.)

Internal implication: $\Gamma, A \vdash B$ iff $\Gamma \vdash A \rightarrow B$

Strong internal implication : $\Gamma, A \vdash \Delta, B$ iff $\Gamma \vdash \Delta, A \rightarrow B$

We show Avron's derivations of operational rules of internal implication for illustration (the case of strong internal implication is similar). The first two are both ways of the definitional equation. The last one is the operational rule $L \rightarrow$

---

[31] Avron also gives the following definitional equation : $A \supset B, \Gamma \vdash \Delta$ iff $\Gamma \vdash \Delta, A$ and $B, \Gamma \vdash \Delta$. Avron calls this implication as combining implication. Here we do not need to discuss this.

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B} \qquad \frac{\Gamma \vdash A \to B}{\Gamma, A \vdash B} \qquad \frac{\Gamma \vdash A \qquad \Gamma', B \vdash C}{\Gamma, \Gamma', A \to B \vdash C}$$

$R \to$ is identical with the first. The third one $L \to$ is derived from the second one as follows. ($A \to B \vdash A \to B$ is an instance of strong reflexivity.)

$$\dfrac{\dfrac{\Gamma \vdash A \qquad \dfrac{A \to B \vdash A \to B}{A \to B, A \vdash B}}{\Gamma, A \to B \vdash B} \qquad \Gamma', B \vdash C}{\Gamma, \Gamma', A \to B \vdash C}$$

In this derivation, unlike previous cases, $\to$ reflect $\vdash$ (yields), but the pattern of reflection is less clear than previous cases. (The reflection is not located on one side. Also, the symbol in the metalanguage to be reflected is not given independently of $\vdash$. $\vdash$ itself is reflected). Nonetheless, in this thesis we take this to be the official solution of the definitional equation for implication for the reasons explained below.

Let us now summarize our basic view (to be somewhat modified later) here, since taking "logical constants" to be characterized by combining structural features of proof systems and reflections of metalinguistic operations may now become a plausible idea. Our basic view of the notion of logical constants in this chapter is as follows. *Logical constants are those constant expressions in the object language which are obtained as the results of solving the definitional equations, used in sequent calculi that are structural variants to one another and that represent abstract consequence relation.*[32] We take

---

[32] It is true that this depends on a particular proof-theoretic framework, but this means that we are committed to the view that those introduced in this way must be the most basic. Also, note again that here we are still identifying $\vdash$ and $\Rightarrow$; hence, a representation may look trivial, but it is not (in general).

the solvability of the definitional equation (in Avron's sense for implication) to be a positive test for logical constant-hood.[33]

Solvability of the definitional equation can be taken as a generalization of the idea of "intrinsic harmony" as Dummett understands it. This is so in the following sense. If we take a double-line rule form in a sequent system in order to formulate the idea involved in the definitional equation, at least in some intuitive sense, we can trivially derive something close to a sequent counterpart of "leveling local peaks." Hence, the intuitive idea of intrinsic harmony, i.e., balance between a ground and a consequence of an assertion, may be implicitly contained in the formulation itself in the sequent format. However, in sequent calculi, such rules are not in an appropriate format in order to prove cut-elimination. To obtain appropriate rules for cut-elimination, solving the equation plays an important role.[34] Still, solving the definitional equation at most shows intrinsic harmony rather than total harmony. Since in sequent calculi proof-theoretic properties of a system are determined by combinations of operational rules and structural rules, solving the definitional equation itself may not imply cut-elimination or conservativeness. On the other hand, Dummett's intrinsic harmony is formulated only for intuitionistic logic (in natural deduction).[35] Hence, the coverage of solvability

---

[33] The works cited so far do not use their argument for reflection in order to give an answer the question of logical constant-hood. However, the connection between Sambin et al. [143] and the issue of logical constant-hood is already made in [26]. See footnote 89 in section 2.4.8.

[34] Note also that all the solutions use only reflexivity and transitivity. No other structural rules are required.

[35] Introduction rules are taken as self-justifying, and his total harmony is based on

of definitional equations in sequent calculi is much broader (covering most substructural logics and more).

b. Are modal operators logical constants?

However, this approach leaves unsolved some potential problem that is pointed out by Sambin et al. The problem is discussed in [143] as follows.

> The peculiarity of implication is that it reflects a link *yields*, that is the turnstile sign $\vdash$ itself, and this is what makes it different from other connectives. So to see that implication follows the same conceptual pattern as all other connectives, a richer meta-language is needed, in which the link *yields* to be reflected appears inside the scope of another link *yield*. In terms of the shorthand notation, also nested occurrences of $\vdash$ must be considered. Then the definitional equation for $\rightarrow$ is simply
>
> for all Γ, $\Gamma \vdash A \rightarrow B$ if and only if $\Gamma \vdash (A \vdash B)$.
>
> It can be solved following the same pattern as other connectives, but it needs two new forms of composition, that is composition of

---

normalization (in a way it can be regarded as a better property especially if it is strong normalization).

> formulae inside two occurrences of $\vdash$. Then one reaches the rules
>
> $$\frac{\Gamma \vdash (A \vdash B)}{\Gamma \vdash A \to B} \qquad\qquad \frac{\Gamma \vdash A \qquad B \vdash \Delta}{A \to B \vdash (\Gamma \vdash \Delta)}$$
>
> which however cannot be expressed in the traditional shape of sequent calculus, where nested occurrences of $\vdash$ are not considered. (p.990 [143])

This becomes a serious problem only if the reflection for $\to$ requires the same standard as other cases of reflection, i.e., the symbol in the metalanguage to be reflected can be identified independently of the main $\vdash$. From a slightly less strict point of view than Sambin et al.'s (e.g., Avron's view), there is no immediate reason why we have to follow this standard when reflecting on $\vdash$. (Sambin et al. themselves decided to choose a weaker version of the definitional equation $A \vdash B$ iff $\vdash A \to B$ so that the formulation does not require nested $\vdash$.)[36]

However, considering in what kind of proof-theoretic framework we can

---

[36] Sambin et al. note that $\Gamma, A \vdash B$ iff $\Gamma \vdash A \to B$ is not derivable in the system B of basic logic (BL). The official rules for $\to$ in their system are as follows.

$$\to\text{-formation } \frac{A \vdash B}{\vdash A \to B} \qquad \to\text{-reflection } \frac{\vdash A \quad B \vdash \Delta}{A \to B \vdash \Delta} \qquad \to\text{-unified } \frac{A \vdash B \qquad C \vdash D}{B \to C \vdash A \to D}$$

The equation is not completely satisfactory because the second rule is not derivable from it (the third one is also added to have the replacement of equivalent formulas), but it satisfies the condition of "visibility," and they claim that these rules can be justified by cut-elimination in B. Hence, their weaker version of the definitional equation barely satisfies their needs.

possibly express such nested $\vdash$, we may be able to gain some further insight about the nature of implications. To explore such a possibility, we will take a look at Došen's higher-level sequents [44] as a candidate of a proof-theoretic framework in which nested $\vdash$ can be expressed. Higher-level sequents have a form such as $(\Gamma \vdash^1 \Delta) \vdash^2 (\Sigma \vdash^1 \Pi)$ (Došen has $\vdash^n$ for any finite $n \geq 1$). Hence, at least naively, it looks like the nested $\vdash$ above. Došen's motivation of using it is to formulate a proof-theoretic framework via the notion of "horizontalization" introduced by Dana Scott. Thus, let us also start explaining Scott's idea of horizontalization.

In [151] Scott discusses how to make a clear metamathematical distinction between the following versions of modus ponens.[37] ($\Longrightarrow$ is Scott's implication.)[38]

$$\text{(MP1)}\ \begin{array}{l}\vdash A \Longrightarrow B\\ \vdash A\\ \vdash B\end{array} \qquad \text{(MP2)}\ A, A \Longrightarrow B \vdash B \qquad \text{(MP2}'\text{)}\ \begin{array}{l}\Gamma \vdash A \Longrightarrow B\\ \Gamma \vdash A\\ \Gamma \vdash B\end{array}$$

Also, though Scott does not discuss the following cases, to make our discussion clear, we add a similar issue of a distinction between two different formulations of some principles (or rules) roughly corresponding to the de-

---

[37] We are focusing on Scott's concept of horizontalization, but, in a series of papers [148], [149], [151], [150], Scott had at least two general motivations. One is to formulate a generalized consequence relation (in 2.2). The second is to apply the idea to modal logics and many-valued logics.

[38] $\vdash$ is given by Scott's consequence (i.e., the same one given in 1 below). Also, note that MP2$'$ is from [122], and this is equivalent to MP2 over Scott's system.

duction theorem.[39]

$$\text{(DT1)}\ \frac{A \vdash B}{\vdash A \Longrightarrow B} \qquad\qquad \text{(DT2)}\ \frac{\Gamma, A \vdash B}{\Gamma \vdash A \Longrightarrow B}$$

Note that modus ponens and the deduction theorem are sometimes mentioned as characteristic features of implications. But still there may be subtleties about particular formulations. Which commonly used logical principles correspond to which versions may be a non-trivial problem.

To make a distinction between MP1 and MP2, Scott introduces a logical system for an implication (we call it SSI) and shows a proposition for SSI. (The language of Scott's SSI has $\Longrightarrow$, $\wedge$, $\vee$, $\top$, $\bot$.)

1. Rules for $\vdash$: reflexivity (R); monotonicity(M); transitivity (T).
2. The rule of confusion $\qquad \text{(C)}\ \dfrac{A \vdash B}{\vdash A \Longrightarrow B}$ (double line)
3. Three rules for implication $\Longrightarrow$.

$(\mathrm{MP}_{\Longrightarrow})\qquad A \Longrightarrow B, (A \Longrightarrow B) \Longrightarrow (C \Longrightarrow D) \vdash C \Longrightarrow D$

$(\mathrm{T}_{\Longrightarrow})\qquad A \Longrightarrow B, C \Longrightarrow D \vdash (B \Longrightarrow C) \Longrightarrow (A \Longrightarrow D)$

---

[39] These things may not be identical with what is usually called the deduction theorem since traditionally the deduction theorem is formulated in deduction from assumptions where assumptions are taken to be sets. R$\rightarrow$ is not, in general, the same as the deduction theorem. In particular, in substructural logics, they may be radically different, due to the gap between what is assumed in a consequence relation and what is assumed in a sequent calculus. However, in this subsection, we consider only cases where set-hood of formulas in consequence relation is assumed (as we will present shortly). Thus, this raises practically no problem. Also, there is an important intermediate case between these two formulations of DT. We will discuss this later, but here it suffices to make the distinction of DT1 and DT2. A case of DT1 goes back to [21].

$(\mathrm{T}_{\wedge,\vee})$ $\quad A \Longrightarrow (B \vee C), (A \wedge B) \Longrightarrow C \vdash A \Longrightarrow C$

4. Rules for $\top$, $\bot$, $\wedge$, and $\vee$. $\quad(\top)\ \vdash \top$; $\quad(\bot)\ \bot \vdash$

$(\wedge)\ A, B \vdash A \wedge B;\quad A \wedge B \vdash A,\ A \wedge B \vdash B$

$(\vee)\ A \vee B \vdash A, B;\quad A \vdash A \vee B,\ B \vdash A \vee B.$

To state the proposition, we need a few definitions.[40] A collection of rules is *normal* if it contains $(\top)$, $(\bot)$, $(\wedge)$, $(\vee)$, and (C). Two collections are *equivalent* if each rule of one is derivable from rules in the other.

Proposition 2.4.2 (Horizontalization) *A normal collection of rules is such that whenever a rule of the vertical form :*

(V)

$$\begin{array}{c} \Gamma_0 \vdash \Delta_0 \\ \Gamma_1 \vdash \Delta_1 \\ \vdots \\ \dfrac{\Gamma_{n-1} \vdash \Delta_{n-1}}{\Gamma_n \vdash \Delta_n} \end{array}$$

is derivable, then the horizontal form :

(H) $\bigwedge \Gamma_0 \Longrightarrow \bigvee \Delta_0, \bigwedge \Gamma_1 \Longrightarrow \bigvee \Delta_1, \ldots, \bigwedge \Gamma_{n-1} \Longrightarrow \bigvee \Delta_{n-1} \vdash \bigwedge \Gamma_n \Longrightarrow \bigvee \Delta_n$

---

[40] Scott has a precise definition of "derivability of a rule," but this is not very different from the commonly known definition of derivability of a rule in a proof system, so we omit it.

is derivable, if and only if it can be replaced by an equivalent normal collection containing aside from (C) only horizontal rules including ($MP_{\Longrightarrow}$)*,* ($T_{\Longrightarrow}$)*, and* ($T_{\wedge\vee}$)*.*

Scott was not interested in proof-theoretic aspects of SSI (SSI is not cut-free) but in clarifying the interaction between the object language and the metalanguage via a sequent-style system. The main point of the proposition can be described as showing how to "reflect" a horizontal line and $\vdash$ in terms of $\vdash$ and $\Longrightarrow$ via the rule called "confusion." (Note that "Confusion" is like the restricted version of the definitional equation.) Although Scott assumes other auxiliary rules in order to formulate horizontalization, Scott showed that the proof does not require MP2. (In other words, the auxiliary rules are weaker than MP2.)

On the other hand, Došen [42] formulate the following double-line rules in his higher-level sequents to capture Scott's insight from his structural view (here we assume set-hood and monotonicity for Došen's system, too. For details of Došen's system including notations, see the appendix).

$$(\mathrm{I})\ \frac{\{A\} \vdash^1 \{B\}}{\overline{\emptyset \vdash^1 \{A \Longrightarrow B\}}}$$

$$(\mathrm{II})\ \frac{\{\{A_1\} \vdash^1 \{B_1\}, \ldots, \{A_{n-1}\} \vdash^1 \{B_{n-1}\}\} \vdash^2 \{\{A_n\} \vdash^1 \{B_n\}\}}{\overline{\emptyset \vdash^2 \{\{A_1 \Longrightarrow B_1, \ldots, A_{n-1} \Longrightarrow B_{n-1}\} \vdash^1 \{A_n \Longrightarrow B_n\}\}}}$$

$$(\mathrm{III})\ \frac{\Gamma + \{A \vdash^1 B\} \vdash^2 \Sigma + \{\Theta \vdash^1 \Xi\}}{\overline{\Gamma \vdash^2 \Sigma + \{\Theta + \{A \Longrightarrow B\} \vdash^1 \Xi\}}}$$

$$(\mathrm{IV})\ \frac{\Gamma + \{A\} \vdash^1 \Sigma + \{B\}}{\overline{\Gamma \vdash^1 \Sigma + \{A \Longrightarrow B\}}}$$

It seems that Scott's question is a legitimate one independently of what kind of implication we obtain in his system and Došen's corresponding system. However, the answer to the question of what these implications are is more than just a matter of curiosity. It turns out that they correspond to the following logics.

SSI = Došen's system with (I) and (II) = weak S4 strict implication[41]

Došen's system with (III) = S4 strict implication (and S5 strict implication)

Došen's system with (IV) = Intuitionistic implication (and classical implication).

Došen's formulation's intuitionistic and classical logics (and S4 and S5) are based on the idea that we have already mentioned in 2.4.1.2. (namely, restricting applications of weakening on the succedent). That is why we mentioned S5 and classical logic in parentheses. Weak S4 (wS4) is a modal logic obtained by $\Box\Box\varphi \to \Box\varphi$ (weak density) to K4 in the traditional formulation of axiomatic system with material implication ($\to$). What Scott formulated turns out to be a strict implication version of wS4.[42]

These are already interesting observations since these essentially show

---

[41] Scott does not give any proof of non-derivability of MP2 in SSI. However, this observation in [42] and the fact that MP2 exactly corresponds to T axiom in strict implication logic [37],[87] imply that MP2 is not derivable in SSI.

[42] wS4 is a minimal normal modal logic to which positive intuitionistic logic can be embedded ([43])

that in order to make sense of the metatheoretical distinction between MP1 and MP2, we may need modal logic.[43] Scott's original motivation was to give a reply against Quine's critique of modal logic claiming that there is a use-mention confusion in C.I. Lewis' formulation modal logics. However, independently of whether Scott's formulation is successful as a reply to Quine's critique, this already seems to give a sufficient raison d'être of modal logic.

Then let us quickly mention what we can say about Sambin et al.'s nested $\vdash$ in view of the above results. We can derive (*) $\Gamma \vdash^2 \{A \vdash^1 B\}$ iff $\Gamma \vdash^2 \{\emptyset \vdash^1 A \Longrightarrow B\}$ in Došen's higher-level sequent calculus by using (III) (see the appendix). (IV) is rather a notational variant of $\Gamma, A \vdash B, \Delta$ iff $\Gamma \vdash A \to B, \Delta$ (with structural assumptions for $\vdash$). In (*), levels never collapse, but Sambin et al.'s lower sequent does not have any higher level. In fact, for most logics (except very weak cases such as B) that Sambin et al. considers, a principle that looks like DT2 (in the form of R$\to$) holds. This means that this holds for most substructural logics. Hence, nested $\vdash$ may give merely a roundabout way to derive R$\to$ for these logics. However, for logic for which DT2 fails, nested $\vdash$ may be needed to introduce $\to$ without collapsing levels. This suggests that Sambin et al.'s nested $\vdash$, as it is, may not exactly correspond to any extant commonly used logic.[44] But it may indirectly show what difference exists between implications in substructural

---

[43] Certainly, this does not show modal logic is "necessary" but modal logic is sufficient to make a distinction. However, it is highly likely that more or less the same principle will be involved even if we use another way of making a distinction.

[44] We may need some change of view to appropriately take a look at how nested $\vdash$ works in order to consider real applications.

logics and implications outside of them.

Why does this matter to us? The eventual goal of the remaining part of the whole subsection 2.4.3.2. is to argue that modal operators in most major normal modal logics are (decent) logical constants based on our conception of logical constants. For this purpose, this is of crucial importance, since we want to show this via first arguing that strict implications corresponding to these modal logics are logical constants by using a method motivated by Scott's idea of (C).

Although we presented Došen's higher-level sequents for formulating these logics, in the official presentation of our argument to show that these strict implications can be decent logical constants, we first keep using traditional sequent calculi and then we move on to a modest (proof-theoretically satisfactory) extension of them.

To explain the reason for this choice, we first formulate some adequacy conditions for a proof-theoretic framework (or proof systems) that can formulate the implications mentioned above.[45]

(1) being able to make a distinction between MP1 and MP2

(2) being able to make a distinction between DT1 and DT2

(3) enjoying cut-admissibility

The current situation can be summarized as follows. Scott's approach to strict implications can cover (1). Došen's approach can cover (1) and

---

[45] Scott's case is only one system, but by using the idea of adding auxiliary rules, his approach can be regarded as formulating logics by using traditional sequent calculi and auxiliary rules.

(2). However, neither of these satisfies the condition (3). (Unlike Scott or Došen, we take cut-admissibility to be one of the most important properties in any proof theoretic framework (see 2.4.6.). Cut-admissibility may not automatically hold in a framework that can characterize logical constants. Thus, it would be ideal if we can find out a proof-theoretic framework that satisfies all the three conditions. However, currently there seems to be no proof-theoretic framework which satisfies all these three properties. Hence, in this thesis, we will be temporarily satisfied with keeping (2) and (3) as our theoretical goal and leave the issue of (1) to future work.[46] To achieve this goal, we introduce two modifications in the traditional sequent calculi.

Let us now discuss the following issues. First, we show how we can generalize Scott's view to other strict implications without using auxiliary rules. This contains the first modification among the two, namely using some formulas in the object language in a context in a sequent calculus. Second, we define modal operators by these strict implications. This is the basis of our claim that modal operators are logical constants, since there is a reason to think that the strict implications are logical constants. Third, we will

---

[46] We conjecture that the framework that we adopt, namely hypersequent calculi, can at least partially handle the condition (1), but we will not go into this direction here. Since the topic is relevant, let us briefly comment on nested sequents (aka deep inference). Nested sequents have been developed to handle some proof-theoretic problems so that we can manipulate formulas deep inside the structure of a formula, but they also have applications in modal logics ([29]). They can be used to formulate a cut-free proof system for a logic that is hard to formulate by a traditional sequent calculus (e.g., modal logic B). The idea of nested sequents may be naturally connected to nested $\vdash$. It must be interesting to explore this connection between nested $\vdash$ and nested sequents. (See the relevant footnote in 2.5).

present cut-free (hyper)sequent calculi for these modal logics. This is the second modification (in fact, an extension) of sequent calculi.

b1. Applications of Scott's method beyond (C)

We will now explore strict implications other than wS4 from a "structural" point of view. In the following discussion of modal logic, we assume monotonicity and set-hood of formulas in our consequence relation in addition to reflexivity and transitivity, unless otherwise mentioned.[47]

As already mentioned, Scott's approach to strict implication based on "confusion" (C) is almost the same as Sambin et al.'s $A \vdash B$ iff $\vdash A \to B$.[48] Then it is a natural question whether, apart from horizontalization, there is a way of formulating traditional strict implications other than wS4 along the line of Avron's derivation of $\to$ rules. Došen already gave one way of formulating S4 strict implication via higher-level sequents ([44]). However, there is another economical way of formulating strict implications by using almost the same definitional equations as Avron's.[49] To present the method, we introduce the following notation. The notation $\Gamma_{\Longrightarrow}$ means all the formulas $\varphi_i (1 \leq i \leq n)$ in $\Gamma = \varphi_1 \ldots \varphi_n$ are of the form $\varphi_i = A_i \Longrightarrow B_i$, namely

---

[47] This is because we mainly discuss normal modal logics here. This *by no means* conceptually exclude the possibility of considering substructural variants of strict implication logics or modal logics. For instance, see [50].

[48] This is so except that (C) assumes set-hood of formulas.

[49] This is our own extension of Sambin et al. and Avron's method to strict implications and modalities.

the outermost logical symbols are all strict implication $\Longrightarrow$. Then we can formulate (C') and (C'') as follows.

$$(C')\,\frac{\Gamma_{\Longrightarrow}, A \vdash B}{\Gamma_{\Longrightarrow} \vdash A \Longrightarrow B}$$

$$(C'')\,\frac{\Gamma_{\Longrightarrow}, A \vdash B, \Delta_{\Longrightarrow}}{\Gamma_{\Longrightarrow} \vdash A \Longrightarrow B, \Delta_{\Longrightarrow}}$$

The idea is to restrict the context in an appropriate manner by using the strict implication symbol in the object language (note that this is the crucial point and our first modification of traditional sequent calculi). Intuitively, this amounts to restricting the contexts to only formulas that reflect $A_i \vdash B_i$ since $\Longrightarrow$ is the result of reflecting $\vdash$ via (C), which is a special case of (C'). The intuitive correspondence between the biconditional (*) derived in Došen's higher-level sequents for S4 implication and our own formulation via (C') (or (C'')) is quite clear if we consider this conceptual origin of $\Longrightarrow$ and the fact that in Γ we have only formulas of the form $\Sigma \vdash^1 \Delta$. In the following, we call SSI′, SSI″ respectively the systems obtained by replacing (C) in SSI with (C'), (C'').

We have several observations here. First, Scott's auxiliary rules $\mathrm{MP}_{\Longrightarrow}$, $\mathrm{T}_{\Longrightarrow}$, and $\mathrm{T}_{\wedge,\vee}$ are all derivable by (C'). Second, the following "operational rules"

$$\mathrm{L}{\Longrightarrow}\,\frac{\Gamma \vdash A \qquad \Gamma', B \vdash C}{\Gamma, \Gamma', A \Longrightarrow B \vdash C} \qquad\qquad \mathrm{R}{\Longrightarrow}\,\frac{\Gamma_{\Longrightarrow}, A \vdash B}{\Gamma_{\Longrightarrow} \vdash A \Longrightarrow B}$$

can be derived in exactly the same way as Avron's derivation of the corresponding rules for internal implication (namely, only via reflexivity and transitivity in the consequence relation). Third, it turns out that these are the operational rules for a traditional (cut-free) sequent calculus for S4 strict implication logic [117]. The multiple-conclusion case gives a sequent calculus for S5 strict implication logic, which is not cut-free, but this can be fixed by using hypersequents (see b3 and Chapter 3). Fourth, (C') derives MP2 (MP2$'$), which implies that SSI$'$ is strictly stronger than SSI. Hence, SSI$'$ is a logic that contains horizontalization and MP2.[50] Fifth, (C), (C'), (C'') look like structural variants except that the occurrences of items in the object language prevent them from being structural variants, according to Došen's criterion. In spite of the fact that the patterns in which they differ from one another is very similar to those of other implications, Došen's criterion of structural deductions states that a system in the deductive-metalanguage is structural when no constant in the object language is essential. The use of strict implications in the context to specify the strict implication rules is essential to formulate them. Hence, according to this criterion, these are *not* "structural" variants. However, it seems very natural to introduce an extension of the view according to which we can take these to be structural variants, due to the striking similarity between strict implications and other

---

[50] MP2 itself corresponds to (T) axiom in strict implication logic [87]. Hence, S4 strict implication logic is a logic of (a strong form of) modus ponens. The whole logic with (C') corresponds to SSI plus MP2, but it is not clear whether $(C')\downarrow$, which is like a restricted deduction theorem, can be directly formulated via SSI + MP2.

implications. We do not intend to replace the traditional notion of schema and structural features. But, presenting technical results (in Chapter 3) that can support for the idea that they can be a kind of structural variants, we make a plea (in 2.4.4.) for slightly loosening the notion used in the criterion of logical constant-hood from purely structural features to semi-structural features.[51]

b2. Introducing modal operators

Putting off an argument showing that the notion of structural variants should be extended, we first present how to introduce modal operators by using strict implications. Once we set up these systems for strict implication logics, introducing modal operators into these systems is straightforward. We can use the definition (1) in [148]. (Incidentally, a negation can be introduced by (2) as if it were a mirror image of a modal operator. For "strict negation," $\neg A = A \Longrightarrow \bot$)

$$(1)\ \Box A = \top \Longrightarrow A. \qquad\qquad (2)\ \neg A = A \rightarrow \bot.$$

One could describe modal operators and negations from a uniform perspective as Restall does in [138]. Most commonly known negation, e.g. de

---

[51] The situation can be seen from a different point of view. Namely, by taking weaker versions of adequacy conditions for a decent implication mentioned above (1. the deduction theorem; 2. modus ponens [122]), we are considering a broader class of implications.

Morgan, intuitionistic, classical, can be introduced via the same method. This view enables us to take most negations as logical constants that reflect $\vdash$. Concerning negations, we are satisfied with this observations and not going deeply into the subject.[52]

To formulate modal logics, it may be better to have material implication (although this is not mandatory). Since we currently assume the same structural conditions as Scott's, we can quickly get it by deriving the operational rules from the definitional equation for strong internal implication.[53]

Based on the definition of $\Box$, we have modal versions (C), (C'), and (C'') corresponding to SSI, SSI', SSI'', called $\mathsf{SSI}^{\Box}$, $\mathsf{SSI}^{\Box\prime}$, and $\mathsf{SSI}^{\Box\prime\prime}$, respectively. (To save space, we present these horizontally. Nothing rests on this notational change.)

$(C^{\Box})\ \vdash A \text{ iff } \vdash \Box A \qquad (C^{\Box\prime})\ \Box\Gamma \vdash A \text{ iff } \Box\Gamma \vdash \Box A$

---

[52] There are quite a few cases of negation that cannot be handled in this way. In some cases, this is because the system itself is more like a hybridized system such as a case in which we have some relevant implication and de Morgan negation. But also there exists a whole hierarchy of "logics" ([20]) including first-degree entailment, Kleene three valued logic, and Priest's logic of paradox whose sequent style formulations lack traditional rules for negation such as follows.

1. Symmetric negation: (1) $A, \Gamma \vdash \Delta$ iff $\Gamma \vdash \Delta, \neg A$; (2) $\Gamma \vdash \Delta, A$ iff $\neg A, \Gamma \vdash \Delta$
2. Asymmetric negation : (1) $A, \Gamma \vdash$ iff $\Gamma \vdash \neg A$; (2) $\Gamma \vdash A$ iff $\neg A, \Gamma \vdash$

These roughly correspond to multiple-conclusion sequents and single-conclusion sequents. To uniformly handle these logics and logics that can be handled in the framework presented in 2.5., we may need a more flexible framework.

[53] Conservativeness is not required for logicality in our view, but it is desirable. Cut-elimination for the entire system implies that adding material implication is a conservative extension of the original system.

$(C^{\Box\prime\prime})$ $\Box\Gamma \vdash A, \Box\Delta$ iff $\Box\Gamma \vdash \Box A, \Box\Delta$

Then the following formulas or rules are derivable in $\mathsf{SSI}^{\Box}$, $\mathsf{SSI}^{\Box\prime}$, and $\mathsf{SSI}^{\Box\prime\prime}$, resp. $\mathsf{SSI}^{\Box}$ (= $\mathsf{wS4}$) proves items in 1.

1. (1) $\Box A, \Box(A \to B) \vdash \Box B$, (2) $\Box A \vdash \Box\Box A$,
(3) $\Box\Box A \vdash \Box A$, (4) $\vdash A / \vdash \Box A$.

2. $\mathsf{SSI}^{\Box\prime}$ proves (1) $\dfrac{\Box\Gamma \vdash A}{\Box\Gamma \vdash \Box A}$ (2) $\dfrac{A, \Gamma \vdash \Delta}{\Box A, \Gamma \vdash \Delta}$

(1′) $\dfrac{\Box\Gamma, \Delta \vdash A}{\Box\Gamma, \Box\Delta \vdash \Box A}$ (2′) $\dfrac{\Gamma, \Box\Delta, \Pi \vdash \Sigma}{\Box\Gamma, \Box\Delta, \Pi \vdash \Sigma}$

3. $\mathsf{SSI}^{\Box\prime\prime}$ proves (1) $\dfrac{\Box\Gamma \vdash A, \Box\Delta}{\Box\Gamma \vdash \Box A, \Box\Delta}$ (2) $\dfrac{A, \Gamma \vdash \Delta}{\Box A, \Gamma \vdash \Delta}$. (1′) $\dfrac{\Box\Gamma, \Theta \vdash A, \Box\Delta}{\Box\Gamma, \Box\Theta \vdash \Box A, \Box\Delta}$

1 is due to Scott [151], but the remark that $\mathsf{SSI}^{\Box}$ and (1),(2),(3),(4) are mutually derivable is in [42]. Scott discusses only (C) (and simply uses other auxiliary rules). However, once we adopt $(C^{\Box\prime})$, it is clear that the double-line rules $(C^{\Box\prime})$ and $(C^{\Box\prime\prime})$ derive the above rules. 2(1) and 3(1) are just the downward rules, and 2(2) and 3(2) can be obtained by the same pattern of derivation as solving the definitional equations for implications (only by using reflexivity and transitivity). Also, (1) and (2) for both cases are the same as modal rules in sequent calculi (we call these Curry-style rules) for cut-free

$\mathsf{S4}$ ([38]) and (not cut-free) $\mathsf{S5}$ ([115], [116]). Hence, we can obtain $\mathsf{S4}$ and $\mathsf{S5}$ from SSI, by defining $\Box$ via $\Longrightarrow$ and $\top$ and by relaxing the "structural" conditions of (C).

By using a slightly more general form of $2(1')$ and $2\ (2')$ that are derivable by $(C^{\Box\prime})$, $2(1)$ and $2(2)$, we can obtain rules for traditional (cut-free) sequent calculi for some weaker modal logics by taking special cases of the rules.[54] The K rule, the D rule, the $\mathsf{4}$ rule and the $\mathsf{KD4}$ rule can be so obtained from $2(1')$ and $2(2')$. Essentially the same observations apply to the rules in $3(1')$. In order to obtain cut-free $\mathsf{K45}$, $\mathsf{KD45}$ rules, we can do the following : (i) modifying A to be Π where Π has at most one formula ($\mathsf{K45}$, $\mathsf{KD45}$) and (ii) allowing $\Pi, \Box\Delta$ to be empty (for $\mathsf{KD45}$)). Hence, sequent calculi for most major normal modal logics that proof-theoretically behave well in traditional sequent calculi (and a few more) can be formulated either (i) by definitional equations or (ii) indirectly by taking special cases of the variants of the rules of 2 or 3.

b3. Cut-free proof systems for modal logics.

Here we give cut-free systems for modal logics. Cut-elimination for modal logics may not be an easy matter. In [148], Dana Scott made the following comment on cut-elimination (in general), "it took me a long time to realize

---

[54] $\Box\Box\varphi \vdash \Box\varphi$ can be obtained in the same way, but we do not treat this case here (see Chapter 3)

that cut is not eliminable – except in very special circumstances" ([148]). We aim for formulating sufficiently many cut-free proof systems for modal logics so that they can be treated as "structural variants" in order to make more convincing our version of (modified) Došen's principle that is to be formulated shortly. This has required generalizations of traditional Gentzen-style sequent calculi. Although we discuss such generalizations more systematically later (2.5), we will be a little ahead of ourselves and show that these logics can be handled as "structural variants" by cut-free so-called "hypersequent" calculi since it is difficult to formulate some well-known modal logics in traditional sequent calculi in a cut-free manner (hypersequents may not be necessary, but they are sufficient for our purpose here).[55]

Let us briefly explain what a hypersequent is. A hypersequent is a multiset (or a set) of finite sequents put together. A hypersequent written as $\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_n \Rightarrow \Delta_n$, and these "|" are read disjunctively. Also, note that we are not using $\vdash$ here anymore. Now we give a "proof system," so we officially use $\Rightarrow$ (Recall some remarks in Section 2). Here are hypersequent calculi for some modal logics discussed above. (In this subsection, we confine ourselves to typical cases in which operational rules are given via the defini-

---

[55] As we discuss later, there are some other modal logics, but they are anyways difficult to formulate by using combinations of basic operational rules that are derivable from the definitional equations and structural rules in traditional sequent calculi. We have some other minor desiderata, too. See 2.5.

Also, let us note that extending sequent calculi to formulate modal logics so that the proof systems fit the structural view of logic is not against the spirit of Došen's original view, since Došen himself extended the framework to higher-level sequents [44]. The difference is the metatheoretic properties achieved by extending the framework. Unlike Došen, we put the priority to cut-elimination (2.4.6.).

tional equation for S4 strict implication and definition of the modal operators and in which we have external modal structural rules. Chapter 3 has a more systematic presentation).

(1) Classical structural and operational rules

1. Operational rules for classical conjunction, disjunction, and falsum are fixed.

2. Basic internal and external structural rules are fixed (see 2.5 or Ch. 3).

(2) L$\Box$ (= T) rule : $\dfrac{G|\varphi, \Gamma \Rightarrow \Delta}{G|\Box\varphi, \Gamma \Rightarrow \Delta}$ R$\Box$ rule: $\dfrac{G|\Box\Gamma \Rightarrow \varphi}{G|\Box\Gamma \Rightarrow \Box\varphi}$.

3. External modal structural rules:

Modal splitting (S5): $\dfrac{G|\Gamma, \Box\Delta \Rightarrow \Pi, \Box\Lambda}{G|\Gamma \Rightarrow \Pi|\Box\Delta \Rightarrow \Box\Lambda}$

S4.2 and S4.3 are given by adding each of the following rules to S4

4.2 : $\dfrac{G|\Box\Gamma, \Box\Delta \Rightarrow}{G|\Box\Gamma \Rightarrow |\Box\Delta \Rightarrow}$ 4.3 : $\dfrac{G|\Gamma_1, \Box\Delta_1 \Rightarrow \Pi_1 \qquad \Gamma_2, \Box\Delta_2 \Rightarrow \Pi_2}{G|\Gamma_1, \Box\Delta_2 \Rightarrow \Pi_1|\Gamma_2, \Box\Delta_1 \Rightarrow \Pi_2}$

The sequent calculus for S5 having the rules introduced in the last subsection is *not* cut-free. Cut-free formulations of S5 (S4.2, S4.3) satisfying the desired properties (2.5) apparently require some extension of sequents such as hypersequents. To formulate these logics, we add the kind of rules called "external modal structural rules." Došen's definition of structural rules, strictly speaking, does not allow them to be called "structural rules," again due to the use of an item in the object language (but in a way different from the case given b2). However, by using this proof-theoretic framework and a slightly loosened criterion (2.4.4) of logical constant-hood, all these modal logics can be naturally viewed as "structural variants" and, hence, uniformly understandable from a kind of structural view combined with the principle of reflection.

At this point, it may be desirable to give some intuitive justification for the use of hypersequents (provided that multiple-conclusion sequents are accepted[56]). Compare single-conclusion hypersequents and multiple-conclusion sequents. For instance, take a classical case. In this case, single-conclusion hypersequents can do almost everything that multiple-conclusion sequents can do (see [97]). Hypersequents handle multiple "inferences" via $\Rightarrow$ put horizontally, but multiple-conclusion cases have only multiple formulas on the succedent. They can be taken as notational variants of multiple-conclusion sequents. From this point of view, concerning the use of the format of a hypersequent, if one accepts multiple-conclusion sequents but rejects hypersequents, then one may be incoherent (unless one provides a clear reason for the distinction). The case of S5 is quite similar to the classical case. We have both a multiple-conclusion sequent calculus (modal rule is multiple-

---

[56] For this topic, see [139]. We do not make a strong commitment to the view, but it provides some basis for intuitive acceptance of multiple-conclusion sequents. See also [153].

conclusion) and a hypersequent calculus for S5, but unlike the case of classical logic multiple-conclusion sequent calculus for S5 is not cut-free, but the hypersequent calculus is cut-free. There seems to be no reason to accept only multiple-conclusion sequent calculus with cut and to reject cut-free hypersequent calculus.

Concerning technical matters of these logics, we put proofs of cut-elimination for these logics in Chapter 3. We cover S4, S4.2, S4.3, S5, and strict implication counterparts of these logics. We also present sequent calculi for some weaker logics for the sake of completeness of presentation.

2.4.4 A structural-reflective view of logic

Here we give a more systematic presentation of our own view of logic, i.e. structural-reflective view of logic, to clarify on what basis we have extended the notion of structural features in proof systems to formulate modal logics. This also summarizes what we have claimed about the extended notion of structures and reflections. First, we give an outline of the view in contrast with Došen's view of structural deductions. Second, we present our modified notions of structural features that fit well with our view. This amounts to using modal operators as something like "structural markers."

The following general points of Došen are agreeable with suitable qualifications on what we mean by "structural". (We call these shared points "structural view.")

1. Logics are characterized by the definitional equations for logical constants and structural features

2. Different logics should be formulated as structural variants

3. The metalanguage is more basic than the object language

We are now raising the issue concerning what we mean by "structural." From Došen's point of view, structural deductions are deductions that do not essentially involve constants in the object language. On the contrary, we argue for an extension of the notion of structural features to that direction to some extent. Perhaps, it is better to introduce the following terminology to make a clear contrast between Došen's case and ours. (Note that we are dealing with only propositional cases.)

1.1. A (structural or operational) rule is purely schematic if at most the outermost logical symbol of the principal formula (in a case of an operational rule) is in the object language, and otherwise there is no use of any constant in the object language.

1.2. A (structural or operational) rule is partially schematic if the outermost logical symbols of formulas that occurs in contexts are in the object language, but, otherwise, the same condition as a purely schematic case holds.

2.1. A rule is purely structural if the rule is formulated purely schematically and does not introduce any new constant in the principal formula.

2.2. A rule is partially structural if the rule is formulated partially schematically and does not introduce any new constant in the principal formula.

3. A rule is structural if it is either purely structural or partially structural.

We also apply these terminologies to "contexts" similarly. According to these definitions, R rules for modal operators and strict implications are partially schematic operational rules (anyways, these are not structural rules).[57]

---

[57] Schroder-Heister [147] formulates a variety of logics by a general inference schema. He separates the two ingredients of logical rules. One is (higher-level) rules, and the other is inference schema. Different logics are formulated by combination of inference schemata and different primitive rules. Inference schemata play the role of structural rules in traditional sequent calculi.

General inference schema : $(\Delta_1 \Rightarrow A_1), \ldots, (\Delta_n \Rightarrow A_n) \Rightarrow D$ $\dfrac{\Gamma, \Delta_1 \vdash A_1, \ldots, \Gamma, \Delta_n \vdash A_n}{\Gamma \vdash D}$.

This means that we can derive $\Gamma \vdash D$ from $\Gamma, \Delta_1 \vdash A_1, \ldots, \Gamma, \Delta_n \vdash A_n$ via the rule specified by $(\Delta_1 \Rightarrow A_1), \ldots, (\Delta_n \Rightarrow A_n) \Rightarrow D$. Here $\Rightarrow$ is not a symbol in the object language. It is a structural symbol expressing the notion of "rule," which is similar to "," or a horizontal line) in traditional sequent calculi. Primitive rules play the role of operational rules. Schroeder-Heister briefly discusses modal logics $\mathsf{S4}$ based on relevant logic (hence, this is $R^{\Box}$). He formulates rules in modal logic by a metalinguistic structural symbol $\triangleright$.

Inference schema: $\dfrac{\Gamma \vdash R}{\Gamma \vdash \triangleright R}\ (\triangleright +)$ provided Γ is modalized $\dfrac{\Gamma \vdash \triangleright R}{\Gamma \vdash R}\ (\triangleright -)$

where "modalized" means that all formulas in Γ is of the form $\Box\varphi$.

Rules: $\triangleright p \Rightarrow \Box p$ $(\Box p, (\triangleright p \Rightarrow r)) \Rightarrow r$ or (equivalently) $\Box p \Rightarrow \triangleright p$

$\triangleright$ is like Wansing's $\bullet$ (see 2.5.2. and the appendix), although $\triangleright$ has much less structural freedom since it covers only $\mathsf{S4}$ modality. Note that even in the inference schema, Schroeder-Heister uses "modalized" contexts, but Schroeder-Heister still call the inference schemata "structural schemata." This use of the term "schematic" may give a support for our use of the term, since modalized contexts are common to Curry-style modal rules and Schroeder-Heister's schema and this is an extant case in which modalized contexts

All of our external structural rules characterizing modal logics and strict implication logics are structural, but not purely structural. Based on these terminologies, we can call Došen's position "purely structural view."

Note, however, there is a difference between allowing modal operators in a context of operational rule (e.g., $R\Box$ for $\mathsf{S4}$) and using modal operators to indicate that some structural rules are to be applied only to modalized formulas. For the former, a conceptual justification is available since the implication was introduced as a reflection of $\vdash$. Intuitively, the contexts consist of "derivable" formulas (recall the discussion in 2.4.3.2.b2); hence, contexts are restricted to a particular form that indicates this. But in the latter case, by applying structural rules, we typically give a stronger structural property for the modalized formulas. The exponential (!) in Girard's linear logic [67] works in such a way. In fact, when Avron first discussed his hypersequent calculus for $\mathsf{S5}$, he explained the idea of modal splitting mentioning Girard's exponential.[58]

> Here only modalized formulae (those which begin with $\Box$ or $\Diamond$) are allowed to be split from a given component. This limitation is close in spirit to the limitation of the structural rules to modalized (or "exponential") formulae in linear logic. (p.6, [18])

---

are called "schemata."

[58] Restall [140] has a different formulation of $\mathsf{S5}$. Modal rules are not the combination of $\mathsf{S4}$ rules and modal structural rules. But formulating structural variants does not seem to be his goal.

We have not explained this use of modal operators in terms of our official view of reflecting $\vdash$. This use of modal operators to formulate structural rules may need further arguments to fully justify its use, which is, unfortunately, beyond the scope of this thesis, although there is no doubt that formulating modal external structural rules makes hypersequents fruitful from a technical point of view since many modal logics can be made structural variants to one another. Let us just point out that these modalities are used to emulate structural rules ([49]). It takes the opposite role of the first case. A restriction on the contexts in the first sense tends to make the deductive power of the logic weaker.

Next, we discuss a view to be combined with Došen's structural view, namely Sambin et al.'s or Avron's view using "principle of reflection" broadly considered to handle modality. We call this "reflective view." Hence, we call our own view "structural-reflective."

Let us recall that we have identified the basic idea of reflective view as "injecting a part of the metalanguage into the object language." Then Sambin et al.'s idea of solving the definitional equation can be one of particular realizations of this general idea. According to this view, logical constants are those in the object language which satisfy the conditions:

(1) introduced by solvable definitional equations;

(2) symbols so introduced are reflections of operations in the metalan-

guage;

(3) symbols and their definitional equations (and its solvability) are invariant throughout significantly many different logical systems.

Then Scott's work on modal logics can be a preceding example with differences in detail. Our extension of Scott's method follows solving the definitional equation (taking Avron's way). In this case, too, the difference from other cases of implication is the restriction of the context via the strict implication symbol itself. Since our extension of Scott's method uses the core idea of reflective view of logic, we claim that modal operators that are introduced this way are logical constants.

2.4.5 Some issues concerning the structural-reflective view

In this subsection, we will discuss some possible problems or objections that may arise concerning the structural-reflective view that we presented in the last subsection.

The argument showing that modal operators introduced in this way are decent logical constants can be summarized as follows. Classical, intuitionistic, and other substructural implications introduced by reflecting $\vdash$ are decent logical constants. Strict implications can be introduced in *almost* the same manner as other implications (since the operational rules derived from the definitional equation is partially schematic). Hence, we have a sufficiently strong reason why they are decent logical constants (unless we are given a good argument showing that strict implications cannot be decent logical constants). Since modal operators can be explicitly defined via strict implications, modal operators can be taken as decent logical constants.

Although this has some plausibility, some possible objections may be raised. We handle two major objections that we can anticipate here.[59] Both of them rest on the issue that (C') and (C'') use the restricted contexts specified by explicitly using strict implications in the object language.

First, one may object to the current approach by pointing out that the use of strict implication symbols in the contexts spoils the methodological purity of structural view. Indeed, if one takes Došen's definition of structural analysis strictly, then the difference between strict implications and other implications (classical and substructural) is quite substantial. There is one explicit objection, and there is another view that is potentially against our view in the literature about this issue.

---

[59] Let us discuss minor points in this footnote. (1) In modal cases, only *ordinary* consequence relation is considered. Only too few structural variants exist for modal cases. (2) Scott's SSI uses those auxiliary rules other than (C) which are not structural rules.

To (1), we reply that there is no reason why we cannot introduce modalities based on weaker consequence relations, although the resulting modal logic may not be normal. The derivation of the operational from (C') and (C'') does not require weakening or contraction. (Substructural systems for S4 strict implication logic may be worth exploring. One system is actually well-known. A fragment of relevant logic E can be easily obtained by dropping weakening (M) from SSI′. For more recent developments, see [91]. For modal substructural logics, see [50].) To (2), our reply is that our motivation of citing Scott's work came from (C) itself. To cover horizontalization and wS4, we may need more general machinery. However, we believe that we can cover sufficiently many structural variants to make our view convincing.

The second objection raises a conceptual issue. One may object that our double-line rules violate Došen's condition that the logical constants introduced by the biconditional cannot occur on both sides of a rule since this looks circular (Recall the discussion on the condition (1) in 2.4.2.) Indeed, the occurrences of $\Longrightarrow$ and $\Box$ in the context makes it impossible for this definitional equation to derive operational rules in such a way that the introduced constant occurs only once in the principal formula (see 2.5.). In Sambin et al.'s term, it is a failure of "explicit definition."[60] We call this problem "circularity in the double-line rules."[61]

We handle the second conceptual objection in 2.4.5.1. and the first methodological objection in 2.4.5.3. In addition, we discuss the scope of the view in the subsection inbetween.

2.4.5.1 A conceptual issue on strict implication and modal logic

Here we handle the potential conceptual problem. Let us first narrow down the source of the problem. (C) has no circularity that we have worried

---

[60] It can also be seen as a violation of "explicitness" of an introduction rule. See Wansing criticism Curry-style modal rules ([173]) (2.5).

[61] Sambin ([142]) discusses the issue of another kind of circularity. He argues that solving the definitional equation itself does not involve circularity. We assume that there exists a proposition expressed, say by $A \otimes B$, and after solving the definitional equations, we can use the equations and formation rules and reflection rules to "define," say, $A \otimes B$. Sambin says, "this is an inductive definition and hence fully predicative." ([142]). The case of a strict implication is more problematic, since it is not clear whether (C') can be taken as a case of inductive definition. The status of (C') according to the theory of inductive definitions is worth investigating.

about. The metalinguistic $\vdash$ is conceptually prior to implication, and the introduction of implication depend on no other occurrences of implication. Also, (C'') can be justified as a "structural" variant of (C'), once we can justify (C'). Hence, we focus on (C'). Note that $L \Longrightarrow$ rule is not problematic since this is the same as L rules for other implications. The mainly problematic case is $R \Longrightarrow$.

It is true that the rule violates the condition (1) and looks circular, but it may not be clear in what precise sense this rule is problematic. To clarify the point, we compare this with intuitionistic $R \rightarrow$, since there is an interesting source arguing that intuitionistic R$\rightarrow$ is circular in a manner different from "circularity in double-line rules" when it comes to its justifying $L \rightarrow$. We argue that there is a sense in which $R \Longrightarrow$ can avoid a circularity attributed to intuitionistic logic. This may be a conceptual advantage of `S4` strict implication.

The issue goes back to Gentzen's remark on the circularity contained in the intuitionistic interpretation of implication.[62] Gentzen explains the issue as follows.

> What does $\mathfrak{A} \supset \mathfrak{B}$ mean? Suppose, for example, that there exists a *proof* in which the proposition $\mathfrak{B}$ is proved on the basis of the assumption $\mathfrak{A}$ by means of inferences that have already been

---

[62] Gödel also pointed out the impredicativity contained in the BHK interpretation of intuitionistic implication [162]. The precise relationship between these cases is not entirely clear to the current author. This is another topic worth investigating.

> recognized as permissible. From this we infer, by $\supset$-introduction: $\mathfrak{A} \supset \mathfrak{B}$. This proposition is merely intended to express the fact that *a proof is available* which permits a proof of the proposition $\mathfrak{B}$ from the proposition $\mathfrak{A}$, once the proposition $\mathfrak{A}$ is proved. The $\supset$-*elimination* is in harmony with this interpretation: here $\mathfrak{B}$ is inferred from $\mathfrak{A}$ and $\mathfrak{A} \supset \mathfrak{B}$; this is in order, since $\mathfrak{A} \supset \mathfrak{B}$ indicates precisely the existence of a proof for $\mathfrak{B}$ in the case where $\mathfrak{A}$ is already proved.
>
> In interpreting $\mathfrak{A} \supset \mathfrak{B}$ in this way, I have presupposed that the available proof of $\mathfrak{B}$ from the assumption $\mathfrak{A}$ contains merely inferences *already recognized as permissible.* On the other hand, such a proof may itself contain other $\supset$-*inferences* and then our interpretation *breaks down.* For, it is *circular* to justify the $\supset$-inferences on the basis of a $\supset$-interpretation which itself already involves the presupposition of the admissibility of *the same* form of inference. The $\supset$-inferences *which occur in the proof* would in that case have to be justified *beforehand*; but this has its difficulties, especially if the assumption $\mathfrak{A}$ has itself the form $\mathfrak{C} \supset \mathfrak{D}$; if this happens, we have actually no proof for $\mathfrak{D}$ from $\mathfrak{C}$ on the basis of which we could ascribe a *sense* to $\mathfrak{C} \supset \mathfrak{D}$. (pp.167, [161])

This circularity Gentzen talks about (we call it "Gentzen's circularity") is concerned with intuitionistic logic. But a solution of the problem is relevant to ours. In order to avoid Gentzen's circularity, Okada [112] proposes two sequent calculi (see also [113]). According to Okada, the circularity arises in justifying the elimination rule via the introduction rule. The problem is that the introduction rule for intuitionistic logic is too strong since the proof of B from A may contain the same form of inference that we want to justify (i.e., modus ponens in which B follows from $A \to B$ and A). If it is so weak that we cannot have such an inference in the proof of B from A, then the justification process would go through without using any circular argument. Hence, he proposes two systems (SI and WLJ) by restricting the R→ rules.[63]

$$\text{R}\!\to \text{ for SI} \quad \frac{A \vdash B}{\vdash A \to B} \qquad\qquad \text{R}\!\to \text{ for WLJ} \quad \frac{\Gamma_{\to}, A \vdash B}{\Gamma_{\to} \vdash A \to B}$$

In SI, contexts for $R \to$ is simply removed. This rule is essentially the same as $(C) \downarrow$ (given some background assumptions). R→ in SI is certainly non-circular in the sense of the circularity in double-line rules, but we need (C') in order to have stronger implications. Hence, we are more interested in the second one, which is to restrict the context of intuitionistic $R \to$ to $\Gamma_{\to}$ (the same as $(C') \downarrow$).

To show that we can avoid the circularity by restricting R→ in WLJ requires further arguments. To do that, Okada defines a constructive semantics

---

[63] We keep using $\vdash$ for the sake of simplicity, although the systems are officially sequent calculi. Also, WLJ is formulated by using traditional structural rules, but we can formulate essentially the same system by our current setting, namely using set-hood, etc.

for logical connectives of WLJ and he claims that this semantics justify all the inference rules of WLJ (except cut) without Gentzen's circularity.

Definition 2.4.3 (Constructive interpretation) [64]

1. $A \wedge B$ is true if A and B have already been recognized as true.

2. $A \vee B$ is true if one of the formulas A or B has been recognized as true.

3. $A \to B$ is true if we have been able to find a WLJ proof of the sequent $A \vdash B$

Definition 2.4.4 (Validity) *A sequent $A_1, A_2, \ldots, A_n \vdash B$ is valid if, whenever all the formulas $A_1, \ldots, A_n$ are recognized as true, then B is also recognized as true.*

Proposition 2.4.5 *Every provable sequent S is valid.*

The proof is based on double induction on $\omega \cdot deg(S) + l(\pi)$, where $deg(S)$ is the maximal number of logical complexity of all formulas in S, and $l(\pi)$ is the length of the proof of π. We focus on the crucial case where the last inference is of the form.

$$\frac{C_1 \to D_1, \ldots, C_n \to D_n, A \vdash B}{C_1 \to D_1, \ldots, C_n \to D_n \vdash A \to B}$$

---

[64] This is different from ordinary inductive definition since the inductive structure used for a proof of soundness uses a double induction where not only the complexity of formulas but length of proof is used. Hence the interpretation of logical constants is intertwined with length of proofs.

Assume $C_1 \to D_1, \ldots, C_n \to D_n$ are already recognized as true. Then according to our interpretation of $\to$, the sequents (†) $\vdash C_1 \to D_1$, $\ldots$, $\vdash C_n \to D_n$ are provable in WLJ. From this and the upper sequent of the last inference, the sequent $A \vdash B$ is provable. Hence, $A \to B$ is recognized as true. Since WLJ enjoys cut-elimination, this constructive semantics can give a justification of all the inference rules in WLJ.

An immediate question must be what the essential difference between WLJ and intuitionistic logic is. One can quickly point out that in the case of intuitionistic logic, (†) cannot be justified in intuitionistic logic since in intuitionistic logic arbitrary formulas are allowed to occur in the context. However, a deeper conceptual difference may not be obvious. The justification breaks down for intuitionistic logic, but the kind of formulas whose justification breaks down in the constructive semantics for WLJ is, for instance, $A \to (B \to A)$. On the other hand, $(A \to B) \to (C \to (A \to B))$ can be justified by the semantics. The difference may not consist in non-circularity but in necessity ([1]) (although the two points may not be independent).

Although we do not have unqualified agreement with Okada's diagnosis,[65] this diagnosis provides a basis for claiming that it is *not* that the $R \Longrightarrow$ for $\mathsf{S4}$ strict implication is *worse* than intuitionistic logic with respect to the issue of Gentzen's circularity, simply because WLJ is essentially the same

---

[65] By defining modal operators from strict implication, we can consider a fragment of $\mathsf{S4}$ within strict implication. Then apparently WLJ may not be entirely free from self-referential realization via the logic of proofs [99]. However, this does not exclude a possibility that WLJ is "less" impredicative than intuitionistic logic. For purposes here, this would be sufficient.

system as our system for S4 strict implication.[66] This does not directly show that the circularity in the double-line rules is harmless, but this shows that we may need more careful analysis about what is really wrong with this form of rule, since the failure of "explicit definition" may have its gain.[67]

Let us note in passing that the role of cut-elimination in Okada's argument suggests that cut-elimination may be *necessary* to show that Genzen's circularity does not arise in (C'), although cut-elimination may not be sufficient to do so, since intuitionistic logic also enjoys cut-elimination (see 2.4.6).

2.4.5.2 The scope of the view

We now discuss the scope of the view. We discuss two issues: 1. the limitation of the method in modal logics; 2. Proof-terms as logical constants.

Our formal method that is adopted based on our structural-reflective view does not cover all major modal logics. There are modal logics that cannot be formulated by starting from (C') and by taking its "structural" variants or special cases. These include : 1. modal logic B (or 5); 2. Gödel-Löb logic (GL).

Apparently, there are no natural ways of obtaining these logics by reflecting $\vdash$ of the consequence relation. This suggests that what are traditionally

---

[66] Cut-admissibility for S4 strict implication is proven in [117], [87], but our semantic proof given for other purposes given in the appendix implies cut-admissibility.

[67] We do not mean to claim that S4 strict implication is better than intuitionistic logic. Again, for purpose, it suffices to show that S4 strict implication makes sense as a logic, as we characterize it now, and $\Longrightarrow$ works as a logical constant.

called modal logics can be classified into two categories. One can be considered as a result of reflecting "$\vdash$," and the other is the one that cannot. Whether our attempt to show that modal operators are logical constants is successful or not, modal logics characterizable by our method constitutes a proof-theoretically natural class of modal logics. Modal logics in the former category may particularly deserve being called "logical," since these are obtained as results of reflecting $\vdash$.[68] From this point of view, considering how hypersequent calculi S4.2 and S4.3 are formulated (obtained from S4 by adding external modal structural rules), proof-theoretically S4.2, S4.3 may be closer to S4 than, say, B.[69] To consider the status of B or GL, recall that our criterion is a positive test, so the "logical constants" cannot be so characterized does not automatically imply that these are *not* logical constants. But we just need another reason why we should call them "logical constants." At this point, probably, it makes sense to introduce some terminology such as "less logical." For instance, consider the case of GL. This has a cut-free sequent calculus but the modal rule is not standard. This may result from the fact that the modality of GL has an (intended) interpretation in arithmetic provability. Some modal logics may reflect on some metatheoretic notions other than a consequence relation $\vdash$ (e.g., tense logics). In these

---

[68] Without mentioning the issue of reflecting $\vdash$, Burgess ([30]) philosophically discusses the problem "which modal logic is the right one for logical necessity?" Burgess argues that S4 is the logic of proof-theoretic logical necessity, which Burgess calls demonstrability logic and that S5 is the logic of model-theoretic logical necessity, which he calls validity logic.

[69] S4.2 and S4.3 have much more applications than B, although we do not cite each of these here.

cases, a phrase "logic of '. . . '", such as logic of '(arithmetic) provability' or logic of 'time,' tends to be used. In this sense, what we have been doing can be described as an attempt to figure out what is a logic of 'logic' since we started with abstract consequence relations and introduced modal operators reflecting these.

We now briefly discuss the status of proof-terms in the Logic of Proofs LP as logical constants since we discuss this "logic" in Chapter 4. Our own results on justification logics presented in Chapter 4 can also be taken as evidence to show that there is a further uniformity in our way of handling modality.[70] However, proof-terms in LP are not modal operators per se. We only claim that proof-terms are like a skeletal structure underlying modal operators that has an intermediate between modal operators and terms in typed combinatory logic (see [5]) for the following reasons. First, rules of logic of proofs are more schematic and less context dependent than modal logics. We need to restrict neither contexts by modal operators nor $R\Box$ to a single-conclusion sequent like S4 case. (In this sense, proof-terms may be closer to traditional logical constants than modal operators.) Second, proof-terms work in a way analogous to terms in typed combinatory logics (in the sense of Curry-Howard isomorphism. See 2.4.6). Terms in Curry-Howard isomorphism are not usually conceived as logical constants, primarily because they never occur inside of a type (= a formula). In this sense, proof-terms in

---

[70] These lead to unexpected consequences in combining logics. Compare hypersequents for S4LPN in Chapter 4 and combined logic S4+L where L stands for S4.2, S4.3, and S5 in Chapter 3.

LP are closer to modal operators than proof-terms in traditional type theories. Proof-terms in LP work like terms in type theory (in the sense of typing à la Curry), but they go inside of a formula. Third, if we try to consider an analogue of a definitional equation for proof-terms, we get something like

$$\text{For some term } t, \qquad \frac{\vec{x} : \Gamma \vdash A}{\vec{x} : \Gamma \vdash t(\vec{x}) : A}$$

The right to left part gives $L : t$ and the left to right part gives a special case of Lifting Lemma ([6]). Combined with the realization theorem ([6]), this can be a basis for the view that proof-terms work as a skeleton of modal operators.[71]

2.4.5.3 Methodological issues

Whether our structural-reflective view is tenable or not essentially depends on the proviso that we introduce the slightly extended notion of schema. Thus, we now give a philosophical defense here, independently of the fact that it technically works. In the literature, there are at least two authors whose positions on the schematic nature of logic in the cases involving modality are different from ours. As we have noted, Hacking [80] excluded modal logic from his notion of logic. Došen [44] takes modal logic to be within logic via purely structural deduction by extending the framework of sequent

---

[71] In spite of this discussion about the nature of proof-terms here, we keep using conventional names (including the word "logic") in referring to systems that we study in Chapter 4 for the sake of simplicity.

calculi to his higher-level sequents. Both of them take their own positions because they try to avoid having some contexts that have modal operators, although their positions and the reasons for their taking them are different from each other. Hence, we take these cases to be substantial disagreements on the methodological issue of whether a violation of purely schematic formulations of logic is allowed or not. Here we try to indirectly defend our extension by arguing against these positions. Essentially our argument is that keeping a purely schematic view of logic has too much cost to pay.

Hacking's understanding of the situation is not very different from ours, for he states, "modal logic is . . . metalogical"(p.310, [80]) due to its non-local features. (Hacking states the reason for this as follows, "To place a restriction on side formulas is to insist that whether a step in a proof is valid depends on the history of a proof and on the forms of the sentences that occur higher up in the proof"(ibid.).) But the appraisal of it differs. Hacking's reason for excluding modal logic is its non-locality due to the restriction of contexts in Curry-style rules, since according to his conception of logic rules have to be local.

We have two reasons for rejecting Hacking's position. First, there is a similarity between modal operators (strict implications) and other logical constants about the way they are introduced. From our point of view, all of them are introduced by reflection (although there is a difference of degree of easiness to do that). This uniformity exceeds differences. Moreover, non-locality of modal rules does not necessarily disturb cut-elimination, although we need some care of modal cases (see Chapter 3). Second, by giving up modal operators, Hacking also needed to give up the standard formulation of the universal quantifier to keep coherence. Hacking's reason for giving up modal operators is its non-locality; however, the universal quantifier is also non-local in a way similar to modal operators, due to its eigenvariable condition. Hence, Hacking gives up the standard formulation of universal quantifier, and Hacking uses ω-rule. However, Sundholm [160] pointed out that there is a methodological incoherence in Hacking's approach. Introducing quantifiers in this way requires the use of a non-constructive principle, whereas Hacking tries to give a demarcation of logic via finitistic cut-elimination.[72] Note that Hacking's main goal of the paper is actually a demarcation of logic and his position is that ramified type theory is logic since it allows "finitistic" proof of cut-

---

[72] Sundholm's argument can be summarized as follows. Hacking requires $\vdash'$ (deducibility in the new extended language with logical constants) has to preserve bivalence and soundness. (1) "Even if the basic prelogical language is constructivist, the bivalence of L′ resulting from the addition of the universal quantifier can be established only non-constructively."(p.165,[160]) To prove bivalence, assume $\forall xAx \notin T$. We have to show that $\forall xAx \in F$. (In the following, At_i is assumed to be bivalent and $At_i \in T$ and $At_i \in F$ are both decidable, for every term t_i.) The statement that $\neg(\forall t_i(At_i \in T))$ iff $\neg\forall t_i(At_i \notin F)$ holds (under the assumption of the decidability and bivalence of A. Here $\in T(F)$ is an abbreviation of "has the value T (F) in a (fixed) model.") The desired conclusion is derived only if $\neg\forall\neg Cx$ to $\exists xCx$ for decidable C, but this is not constructively valid. (2) Similarly, under the same assumptions on T, F, and A, the soundness of the ω-rule with respect to logical consequence cannot be established without use of the constructively invalid distribution principle $\forall x(Bx \vee C) \to (\forall xBx \vee C)$. Sundholm claims that this shows a methodological incoherence, since Hacking does not show the reason for "restricting the methods of proof in some of the preservation results, e.g., the elimination theorem, but not in others, e.g., soundness."(p.166, [160]) Indeed, Hacking requires that cut-elimination be finitistically provable in his demarcation of logic, but bivalence and soundness can be proven only by assuming not even constructive principles. Sundholm takes finitistic provability to be more restrictive than constructive provability. Hence, the incoherence follows.

elimination. Although we do not discuss quantifiers as subject matters here, we would avoid this incoherence, keeping both the standard formulation of universal quantifier and modal operators as logical constants and accepting non-local features of both if we take into consideration quantifiers. This gives an indirect support for the claim that modal operators are logical constants.

On the other hand, Došen takes modal operators as logical constants, but in order to keep the purely structural view, Došen needed to extend the proof-theoretic framework.[73] His higher-level sequents are sufficient to formulate rules for modal operators without restricting contexts by modalized formulas. However, using higher-level sequents is not without cost. First, the system becomes very complicated and not easy to handle (see 2.5.). Došen's approach to modal logic covers only S4 and S5. Second, cut is not eliminable in general (in any level n, $n \geq 2$ for modal logics). In our view, cut-elimination may not be necessary for all logics, but as we discuss shortly, it is required at least for well-behaved logics. It has the first priority among desirable proof-theoretic properties. (The reasons will be explained in the next subsection.) Thus, giving up cut-elimination and keeping the purely structural (schematic) feature of the proof system is not our option.[74]

These arguments may have only limited consequences since there may

---

[73] It is probably worth noting that in his technical paper [50], Došen uses Curry-style modal rules.

[74] It may look odd to compare the standard of logicality and a feature within logic. However, this is a choice of having a higher standard of logicality and a lower standard for a good proof system or taking the opposite combination. The choice should not be made only by the first point.

well be some other approaches that do not share the problems of Hacking or Došen. (Some extension of traditional sequents seems inevitable to keep purely structural features, but it is not obvious that this implies failure of cut-elimination.) However, our arguments must have given some indirect support for our view.

2.4.6 Cut-elimination as completeness

Our view takes cut-elimination seriously. However, our argument has so far only highlighted the importance of cut-elimination but it has not provided any reason why cut-elimination is so important. Nonetheless, the package of our views cannot be completed without arguing for the conceptual value of cut-elimination since our structural-reflective view rests on cut-elimination in the sense that we think that "good" proof systems have to enjoy cut-elimination.

The status of cut as a structural rule is different from other structural rules, and Došen's structural view does not say much about it. Cut-elimination fails in Došen's higher-level sequent calculi [44], but apparently he does not think of it as a serious defect of the systems. Scott's point of using sequent calculi is that they are useful to express the interaction between the object language and the metalanguage. This feature does not imply cut-elimination. Recall also that Hacking and Kremer claim that cut-elimination is important since it implies conservativeness. We do not necessarily disagree with these, but here we focus on other aspects of cut-elimination. We argue that cut-elimination has its intrinsic value, not only its useful applications.

We give three conceptual reasons why cut-elimination has an intrinsic interest. 1) Cut-elimination can be taken as justification of deduction. 2) Solving equations uses cut; hence, failure of cut spoils the whole point of what is called "the separation property" (see 2.5). 3) Cut-elimination is a kind of completeness.

First, recall Okada's solution of Gentzen's circularity essentially appeals to cut-elimination. Okada calls his constructive semantics "Gentzen semantics" ([113]), probably based on the fact that the perspicuous structure of a cut-free proof is useful to see how the meaning of the statement is related to the structure of a cut-free proof. From this perspective, there is a certain justificatory role in cut-elimination. This theme has been extensively explored in philosophical works on the significance of normalization theorem for intuitionistic logic in natural deduction.[75]

> In the course of my semantical explanation of the elimination rule for implication, I have performed certain transformations which are very much like an implication reduction in the sense of Prawitz. Indeed, I have explained the semantical role of this syntactical transformation. The place where it belongs in the meaning theory is precisely in the semantic explanation, or justi-

---

[75] The process of transforming (possibly indirect) proofs into direct proofs can be taken to be a justification of deduction. The view is presented in [54]. Also, this point is closely related to the issue of meaning, but we defer the issue to the next section.

> fication, of the elimination rule for implication. (p. 37, [101])

This observation can naturally be applied to its counterpart in cut-elimination. Okada's result can be understood the viewpoint of justification of deduction..[76] This theme (but concerning normalization theorem) was touched upon in the section 2.3.(II), but there justification of deduction is given with some revisionistic purpose. We are not taking the issue of revisionism in this thesis and we are not necessarily denying that intutionistic claim that normalization theorem makes a difference between constructive logic and classical logic.[77] However, cut-elimination holds for many more logics. On this basis, cut-elimination can be useful to emphasize a non-revisionistic aspect of a justificatory role that a proof-theoretic procedure can have.

To state the second and third points, it is helpful to cite one passage from Sambin et al [143]. According to them, even in a system that does not have contexts on the antecedent in R→, cut-elimination is crucial. ("composition" is their cut.[78])

---

[76] Interestingly, this is exactly the point about which Gentzen mentions a circularity. Since intuitionistic logic enjoys normalization, it is not entirely clear in what sense Okada's solution is better than intuitionistic logic. But this does not affect our point. Okada's constructive semantics is satisfactory provided that we can make the two points: (1) it is at least not worse than intuitionistic logic concerning Gentzen's circularity; (2) cut-elimination as a tool for justification of deduction works approximately in the same way as normalization.

[77] Even in intuitionistic case, it is not easy to establish the exact correspondence between normalization and cut-elimination ([179], [169]). However, our view does not rest on this.

[78] The difference of the two formulations of cut does not affect our point here.

> If we had added cut as a formal rule, we would not have accomplished our task, which was to characterize each connective by its definitional equation, that is exclusively by the rules directly concerning it. In fact, as we already noticed, composition contains implicit information on all connectives, since it says that all derivations can be composed. Using composition up to now has been a sort of *desideratum*. Almost all the arguments used to solve definitional equations involve composition, and they would have little value if composition were not valid. In this sense, cut-elimination is not an option. (p. 994, [143])

Sambin et al. essentially talk about the separation property (roughly, this means that each operational rule governs only one connective separately from others) at the first part. They say that this property will lose its point if composition is not valid. In fact, it should be not only valid but admissible to keep the point of separation. Cut-elimination does the job. This point is not often made since operational rules are usually given in the first place. Elaborating the understanding of operational rules as derivative to the definitional equation makes clear a potential role of cut.[79] Note that this point is not implied by the subformula property. To show that the

---

[79] Schroeder-Heister [145] has a different view of this matter, who uses what is called "definitional reflection" (a construction similar to clausal definitions) to talk about the definitional equation. This makes it possible to formulate rules in BL without using cut at all. However, he does not discuss implications in the paper.

subformula property holds, it suffices to use analytic cut. However, this conceptual point may be spoiled if we use analytic cut.

The third point is close to Sambin et al.'s view that takes cut as setting desiderata but slightly different from it. In his recent exegetical work on Gentzen's text [60], whose main point is to make clear the reason why Gentzen almost never mentions semantic completeness, Franks interprets cut-elimination as immanent matching of the synthesis and the analysis within one sequent calculus.

> Because the notion of logical consequence appears again in this exact form in the immanent features of the calculus LK, the question of the completeness of that logical system was not for Gentzen a question about how the system corresponds with something beyond itself, but a question about the ability of its analytic fragment to keep pace with its internal consequence relation.(p.376, [60])

The analytic methods should fully represent what is contained in logical consequence. This view, therefore, takes a failure of cut-elimination to be a kind of incompleteness in the framework of sequent calculus. In this sense, the conceptual priority is put on cut-elimination itself to the subformula property. The subformula property itself is important since it puts some limit on the combinatorial possibilities of constructing a proof, but it is more important if it is obtained via cut-elimination.[80]

Note that our view is not exactly the same as Franks' interpretation of cut-elimination, since we do not take logical consequence as immanent in sequent calculi, i.e. logical consequence is completely determined by $\Rightarrow$ in sequent calculi. We take abstract consequence relation as more fundamental. (Our view is that $\Rightarrow$ and $\vdash$ are different). However, this difference is not as substantial as it might appear. The abstract consequence relations talk about no particular object language. For a particular logical system in an object language, we take a cut-free representation of the consequence relation to be the only fully satisfactory one, for the same reason as Sambin et al.'s, since logical constants are specified by the definitional equations whose solutions require transitivity (cut).

---

[80] Girard's comment on cut-free proofs and the subformula property can taken to be a support for our general view here, although the cited result is stated for a much stronger logic. To discuss the issue, Girard introduces a logic hierarchy in second-order logic, Σ^n, Π^n, counting the number of second-order quantifier alternations. (This is not the arithmetical hierarchy. For instance, a numerical quantifier in a Σ^0_1 sentence can be translated into Π^1 since using second-order quantifiers, "x is a natural number" can be expressed by second-order sentence without explicitly mentioning the set $\mathbb{N}$. So, Σ^0_1 corresponds to Π^1.) Then Girard's following statement suggests another view of cut-elimination as completeness: "Now if we formulate logic in sequent calculus, we discover that the subformula property holds for *the same class* Π^1, and fails outside. What does it mean? If we consider cut-free proofs, then all possible proofs are already there, there is no way to produce new ones. In other terms, the calculus is complete - nothing is missing. Observe that this completeness is does not refer to any sort of model, it is an internal property of syntax. Such a property cannot be an accident, it should be given its real place, the first: **The subformula property is the actual completeness**." (p.139, [69])

2.4.7 Proof-theoretic semantics and the operational meaning of logical constants

In the following two subsections, we will discuss what kind of consequences we can draw from our structural-reflective view of logic concerning some philosophical issues summarized in 2.3. In this subsection, we clarify the relationship between Došen's structural analysis (including our extended version) and proof-theoretic semantics. We first take a look at Došen's discussions on meaning, and we briefly survey the idea of proof-theoretic semantics. Then, we argue that at least some elements of meaning can be given to logical constants via a version of proof-theoretic semantics in a manner coherent with Došen's structural analysis. About the issue of meaning from the viewpoint of his structural analysis, Došen comments as follows. ((2), (3) are the conditions in 2.4.2.)

> Though an analysis doesn't give the meaning of an expression, as an explicit definition would. It follows from condition (2) and (3) that an analysis is very closely tied to the meaning of the expression analyzed, and that the analytic equivalence should presumably be a kind of relevant equivalence. But this presumed relevant equivalence need not yield an equivalence as strong as identity of meaning, or propositional identity.(p.372, [45])

The condition (2) contains the phrase “analytically true in L,” which is further explained as “true in virtue of the meaning of the expressions in L.” Apparently, Došen has no explicit account of how we should understand the notion of the meaning of an expression in L. One may wonder whether this is satisfactory or not. However, Došen distinguishes his own program of analyzing logical constants and his thesis [I] from others’ program of “defining logical constants by syntactical means”([45]), which appears to refer to Dummett-Prawitz’s approach, i.e., “proof-theoretic semantic” summarized in 2.3 (II). Došen continues as follows.

> First, the main goal of this program is to show that the meaning of logical constants can be given syntactically, whereas our analyses are neutral with respect to this claim, and are equally compatible with the view that the meaning of logical constants is to be given in a more conventional semantical framework. (p.377, [45])

From this passage, one can see that Došen deliberately left open this issue of meaning. He seems to claim that his structural view of logic can be combined with *any reasonable account of the meaning of logical constants.*

Došen explains the reason for his not giving the meaning of logical constants syntactically as follows.

> Second, the search for a criterion for being a logical constant does not always have a very important place in this program. If the problem of finding this criterion is considered at all, it is usually taken that logical constants will be just those expressions whose meaning can be given by syntactic means, which makes the search for a criterion dependent upon the main goal of the program. Thesis [I] is an attempt to formulate such a criterion without tying it to a thesis on the meaning of logical constants. (p.378, [45])

According to Došen, his view is neutral with respect to what kind of semantics we consider for proof systems based on his structural view. However, he also comments that under the program of giving the meaning of logical constants by syntactical means, searching for a criterion of logical constants is problematic since it depends on the goal of the program. Then his view seems rather against the program, although he states that his analysis of logical constants and thesis [I] are "congenial to this program"(ibid.) This is so unless we provide that view of the meaning of logical constants which makes it clear that Došen's structural view and proof-theoretic semantics are actually compatible. In particular, we need a criterion of logical constanthood that does not depend on whether the meaning of logical constants is given syntactically or not. Apparently, discussions about what kind of view of meaning can play this role are missing from Došen's work. So, we are now trying to give one.

Došen's point does not automatically exclude the possibility that after delineating a reasonable class of logical constants, a characterization of the meaning of logical constants can be (partially) given via operational rules through the definitional equations. Došen says, "it is usually taken," not "it has to be taken." We claim that the use of the definitional equations for this purpose can be based on syntactic considerations prior to the observations about its relevance to meaning, but still some element of the meaning of a logical constant can eventually be given by the definitional equation.

To properly understand which elements of meaning can be given to logical constants by proof-theoretic semantics in a way coherent with Došen's structural analysis, we take a closer look at some background of proof-theoretic semantics.

In an ordinary sense of the word "semantics," semantics is a study of the relationship between linguistic expression and extra-linguistic entities. Classically, such extra linguistic entities as elements in an non-empty set (domain), subsets of the domain, etc. in a structure (or in a model) are associated with linguistic expressions, and we define a satisfaction relation based on them. This is how we give a model-theoretic semantics to a formal language. This semantics is usually identified with so-called truth conditional semantics since the key notion in this semantics is truth in a model.

On the other hand, there are some other traditions of semantics that have much more constructive inclination. As a general name for the other approaches, we can use "proof-conditional semantics." Instead of using the notion of truth or truth in a model, we use the notion of proofs in order to consider the meaning of a linguistic expression in a language. One such approach is Brouwer-Heyting-Kolmogorov interpretation ([168], [6]). This is a semantics in which we associate "informal" proofs to linguistic expressions in the language inductively. This approach was motivated by Brouwer's intuitionistic conception of mathematics, and hence the notion of "proofs" used here is usually taken as a kind of mental construction. A problem of this approach is that it is difficult to provide a mathematically precise theory of construction. (Theories given by Kreisel [93] and Goodman [72] turned out to be inconsistent (see [176]).) There is a different approach called Curry-Howard isomorphism (proposition-as-types) [157], which came from the study of constructive type theory. This view identifies types in a constructive type theory as propositions and typed terms as proofs. This view has a strong support from the observation that normalization procedure for formal proofs in natural deduction corresponds to β-conversion in typed λ-calculi. Artemov's logic of proofs and its provability semantics [6] are yet another approach that gives a way of understanding BHK interpretation in a mathematically precise and $\mathsf{S4}$-compliant way (based on classical logic, which means that this provides a way of understanding constructive reasoning from a classical point of view).[81] BHK, Curry-Howard, Artemov's provability semantics are all proof-conditional semantics in the sense that the meaning of a formula is given by specifying under what conditions a formula has a proof. They give the meaning of formulas by assigning objects called "proofs" (or proof-objects) to formulas.

Among the semantics that can be called "proof-conditional semantics," there is a more radical one which does not require one to assign any object to an expression. The semantics is based on Wittgenstein's slogan "meaning is use." Instead of assigning extra-linguistic entities to expressions, this approach considers that the way we use an expression itself constitutes the meaning of an expression. Here we use the phrase "proof-theoretic semantics" to stand for this particular approach ([90]) (The ones given 2.3(II) are typical cases).

From the model-theoretic point of view, which completely separate semantics and syntax, it may sound odd since from that point of view it is absurd that the meaning of an expression is given syntactically. But if we take the slogan "meaning is use" seriously, the connection between rules in a proof system and the meaning of statements may become more plausible.

---

[81] However, the approach may be neutral with respect to the underlying logic, since, although Curry-Howard isomorphism for classical logic is not easy to formulate, intuitionistic logic of proofs that is arithmetically sound (but not complete) can be naturally formulated ([5]). There exists a decent formulation of intuitionistic logic of proofs that is arithmetically complete ([3], [40]). But these systems may go a little deeper than the conventional proof conditional semantics, so it is not easy to have a simple comparison with traditional Curry-Howard isomorphism.

This can be seen in Prawitz's following comment.

> Gentzen considers besides introductions certain specific inferences that he calls eliminations. We cannot expect these eliminations to be derivable from the introductions in the ordinary sense of being derived inference rules in the system given by the introduction rules. Instead, we have to show that they can be justified *in some semantic way*, which is to say that they can shown to be valid in view of the meaning of the sentences involved. (The emphasis is by us.) (p.510, [129])

This quote from Prawitz can tell us why a *semantic* issue is involved here.

Founders of proof-theoretic semantics have used natural deduction for intuitionistic logic (where structural conditions are not isolated) to formulate their semantic thesis, but some proof-theoretic semanticists have tried to use the framework of sequent calculi to formulate their proof-theoretic semantics, partly because those theorists are interested in substructural logics in general rather than supporting only intuitionistic logic. Wansing ([173], [175]), Schroeder-Heister [144], Paoli ([119],[120]), are among those theorists. In sequent calculi, it is natural to make a distinction between the contribution of structural rules and that of operational rules to the meaning of logical constants. Sambin et al. essentially endorse the basic idea of proof-theoretic semantics as follows (although this endorsement should be somewhat qualified due to an issue to be discussed shortly).

> "One of the main principles of proof theory, put forward by Gentzen and clarified mainly by D. Prawitz, is that the meaning of a connective is determined by rules dealing exclusively with it. This discovery is manifested technically in the theorems on normalization of derivations. One of the principles of contemporary proof theory, promoted by Girard, is that a careful control of structural rules of weakening and contraction permits a finer analysis of the structure of derivations. Basic logic pushes both such principles to their ultimate consequences." (p.980, [143])

Among them, Paoli [119] is most explicit about the conceptual distinction between two notions of meaning. Paoli tries to argue for the existence of two different kinds of meaning that can be specified by using a suitable proof theoretic framework. One of them is an *operational meaning*, "whose comprehension amounts to knowing how to use c in inference processes; it is fully specified by the introduction rules for c".The other is a *global meaning* in the calculus. This is 'specified by the class of the system's theorems (provable sequents) containing c'.

Avron has a similar notion, based on consequence relations.[82]

> The present paper suggests an alternative method of taking rules as defining the meaning of connectives. It had the following two main properties:
>
> - The meaning of a connective is always something which is relative to some consequence relation.
> - What defines a connective is not a set of "introduction rules" but a single rule which is reversible.
>
> The reversible rules which define a connective might introduce it in either the succedent or the antecedent of a sequent. In the first case it usually corresponds to an "introduction" rules of N.D., in the second – to an elimination rule. . . . ([16])

According to Avron, the meaning of logical constants can only be given relative to abstract consequence relation. Avron does not seem to isolate the notion of operational meaning, but this relativity imply that an operational meaning is a special case of a global meaning, where the abstract consequence relation is minimal.[83]

---

[82] At this point, it must be clear that we do not want to take the word "definition" literally.

[83] A very similar idea seems to be partly shared by Tennant [165] and Wansing in [172]. Tennant states, "what I am suggesting is that there is an intrinsic meaning to conjunc-

One of Paoli's motivations in introducing the distinction between these two notions of meaning is to give a reply to Quine's critique against deviant logics that all non-classical logics change the meaning of logical constants (e.g., negation). Against this, Paoli replies by stating, "the operational meaning of negation is the same across a wide range of calculi" (p.539, [119]). Although we are not quite sure of how well this reply to Quine works, this feature of the operational meaning may have motivations independent of it. The operational rules are very stable throughout different logics (so the operational meaning must be the minimal common core of these different logics), and the forms of these rules are like sequent-versions of introduction and elimination rules in natural deduction systems that proof-theoretic semanticists have taken to be paradigmatic cases. In addition, from our point of view, these rules are all ultimately derived from the definitional equations in a uniform way. Hence, our characterization of logical constants through definitional equations may fit well with this idea of isolating the meaning associated with operational rules. If any notion of meaning specified proof-theoretically that is coherent with Došen's structural view, it must be the operational meaning.

---

tion, for example, that is invariant across minimal, intuitionistic, and classical logic" (p. 97,[165]). "The correct consequence relation, insofar as it should arise solely from the meaning of the logical constants, is, naturally, the least relation with respect to which the Harmony (global harmony) of the rules governing those constants can be sustained"(ibid.). Also, Wansing says, "Gentzen-style proof theory is usually associated with a certain philosophy of meaning. The idea is that the schematic introduction rules for any n-ary connective $\sharp$, *together with a set of structural assumptions*, *specify the meaning of* $\sharp$, which has certain consequences for the format of rules" ([172]) (emphasis by the present author).

Indeed, a characterization of the operational meaning does not raise the problem that Došen discusses when he mentions the difference between his own program and Dummett-Prawitz's program. Došen thinks that searching for criteria of logical constant-hood is problematic if one gives the meaning of logical constants by syntactic means, because one "usually" identifies logical constants as constants whose meaning can be given by syntactic means and so the searching for the criterion depends on the main goal of the program. Hence, pursuing these two things at the same time is incompatible. However, our method can doubly avoid the danger of pursuing these two incompatible aims. First, our syntactic criteria may not rest on the issue of whether or not the meaning is given by such syntactic means. Second, the meaning of logical constants may not be purely syntactically given in our method.

Let us comment on the first case. If we identify the criteria of logical constant-hood by a purely syntactic ground that does not rest on their own syntactic specification of the meaning of logical constants, then it does not necessarily depend on the main goal of the proof-theoretic semantics. In our case, this can be given by solvability of the definitional equation, since this was originally motivated by the distinction between the structural part and the operational part of sequent calculi, not by any direct concern of meaning. We can say that the meaning itself is ultimately given by the definitional equation. However, it is not that any definitional equation is solvable, and the criterion of solvability is a purely syntactic matter. Hence, this criterion is at least not patently identified with the observation that these are the expressions whose meaning is given by syntactic means.

Also, our characterization of the meaning of logical constants does not have to be purely syntactic. Giving such a criterion of a logical constant that does not depend on the goal of proof-theoretic semantics may be possible because our solving definitional equations is based on a prior understanding of the metalinguistic link in the deductive-metalanguage, since whether or not it is possible to solve definitional equations is not determined by stipulation but has to be proven. (Note that Sambin et al. and Došen both agreed to the priority of the metalanguage.)[84] This feature makes their approach to the meaning of logical constants more *modest* than theorists in proof-theoretic semantics in intuitionistic tradition, who takes introduction rules to be stipulate or self-justifying[85]. Although Sambin et al. and Došen endorse or are at least sympathetic to proof-theoretic semantics, this point must be a major difference between them and Dummett-Prawitz style proof-theoretic semanticists. It seems more appropriate that Sambin et al. agree to the view that the meaning of logical constants are *represented* or made explicit[86] (rather than "determined") via operational rules.[87]

---

[84] Recall that the only metalinguistic resources that we need are the ones that Sambin et al. use as the metalinguistic links or corresponding to Avron's minimal abstract consequence relation.

[85] See [159] about the distinction between two different views in proof-theoretic semantics. One considers introduction rules as self-justifying, and the other takes "meaning-first" approach. How Sambin et al.'s characterization of the meaning of logical constants via the principle of reflection is related to Dummett's critique against a "modest theory of meaning" should be further investigated.

[86] We are not committed to Brandom's view [28], but this phrase summarizes our point well.

[87] Let us briefly comment on proof-theoretic semantics for modal operators. In a way, our introduction of modal operators via the definitional equations and the derivations of

2.4.8 Conservativeness, uniqueness and harmony revisited

From the viewpoint that we have discussed so far, we revisit the notions of conservativeness, uniqueness, and harmony discussed by Dummett and others. In particular, we discuss Belnap's aforementioned argument against tonk, since while Dummett argues for intuitionistic logic, the scope of Belnap's argument is more general. Let us first summarize Belnap's argument.

Belnap's aim is to argue that there is a way of avoiding introducing a tonk-like connective even if one holds a view according to which the meaning of a logical constant is given by specifying its inferential role. Belnap takes such an introduction of a logical constant as a kind of definition. Since it is a definition, Belnap requires the following conditions for logical constanthood. First, logical constants have to be introduced relative to a consequence relation (Belnap considers only "ordinary" ones). Second, the introduction of logical constants has to satisfy the two conditions that are usually required

---

operational rules are already included in proof-theoretic semantics for modal operators, concerning their operational meaning. (In this sense, our approach is clearest in modal logics including S4.) But we leave to future development more detailed discussions on the meaning of modal operators.

Also, let us mention that proof-theoretic semantics for modalities has been proposed by some people. For instance, Read [133] uses natural deduction with prefixes for possible world. Read claims that "possible world semantics" works as a metaphor. However, it is not clear what he means by "metaphor." As we discuss in 2.5, we think that prefixes carry too much semantic information. A purely proof-theoretic method can cover only limited variety of modal logics, but it is important to make it clear what we can cover by using which method. From our point of view, it is not obvious that S4 and B should be treated by the same proof theoretic method. Such a method is acceptable without any hesitation only if they formulated as structural variants.

for a definition: existence and uniqueness. In Belnap's discussion, these amount to conservativeness and uniqueness (2.3).[88]

Let us state our overall views of these two concepts. The notion of conservativeness is sensitive to its formulation. Due to this sensitivity, it is not necessarily a good idea to put too much philosophical significance on conservativeness. Requiring only uniqueness for any logical constant may yield undesirable consequences due to the "duality" of conservativeness and uniqueness that we will discuss shortly.

We start from the notion of conservativeness. Dummett's initial motivation for discussing the issue of "conservativeness" came from an epistemological concern about the role of deductive reasoning [52]. Against Wittgenstein's idea that deductive reasoning modifies the meaning of the words in the premise, Dummett argues as follows.

> When an expression, including a logical constant, is introduced into the language, the rules for its use should determine its meaning, but its introduction should not be allowed to affect the meanings of sentences already in the language. If, by its means, it becomes possible for the first time to derive certain such sentences from other such sentences, then either their meanings have

---

[88] Hacking requires conservativeness and eliminability based on the same reason. Incidentally, uniqueness is the same condition as the one that Smiley [155] called "functional dependence," but he does not require functional dependence for a condition of being a logical constant.

> changed, or those meanings were not, after all, fully determined by the use made of them. In either case, it will not be true that such a derivation demonstrates that the conclusion holds good according to previously acknowledged criteria. (p.220, [53])

The technical notion of conservativeness (defined in 2.3.), attributed to Post, had been known independently of this issue. However, in philosophy of logic, the notion was originally used in Belnap's reply to Prior's tonk argument ([23]), and Dummett connected with it the issue of epistemic significance. Dummett adopts conservativeness as the underlying idea of "total harmony," and (although Belnap did not have any particularly epistemic concern) Dummett uses the notion of conservativeness to critically examine the epistemic status of classical logic.

The formulation of conservativeness by Belnap [23], cited in 2.3., is standard. However, Belnap's solution to the tonk problem appealing to conservativeness may not be entirely satisfactory. We give two sorts of comments on Belnap's solution. The first ones are general. The second ones are directed to the role that the notion of conservativeness plays in Dummett's discussion of proof-theoretic semantics.

In both cases, the bases of our argument are as follows. First, the tonk issue seems to be primarily concerned with a constant symbol. But conservativeness is a global property of a logical system itself. Having a local control on the constant symbol must be more desirable. Second, the notion of conservativeness depends on contexts, i.e., on both the object language and formulations of proof systems.

Let us start discussing the last point first (and go in the reverse order). The relation between what Dummett apparently wants to claim about his epistemic concern and the technical notion of conservativeness is not obvious, because it is unclear whether or not the kind of meaning distortion Dummett is worried about can be removed by the requirement of conservativeness. We claim that conservativeness, as Dummett wants to use it, is not sufficient for Dummett's requirement of faithfulness to the original meaning in the way Dummett needs it. Necessity is more promising since if a case is not conservative in the technical sense, then it is hopeless to claim that the meaning is undistorted. But let us focus on sufficiency here. The situation is more complicated than it initially looks.

The notion of conservativeness allows various formulations, but no technical concept of conservativeness exactly fits in the purposes of Dummett. Among different formulations of conservativeness, the one cited in 2.3. is an intermediate one. We can formulate a weaker ([165]) or a stronger ([65]) notion of conservativeness as follows. Here we call the former "weak conservativeness." Let $\vdash_{At}$ be a single-conclusion ordinary consequence relation only on atomic sentences of the object language. Suppose logical operators are introduced into the language and that new rules of inference are adopted to govern them. Then $\vdash_{At}$ is extended to a consequence relation $\vdash$ for the whole language. Then these rules conservatively extend the atomic fragment if and only if for all atomic sentences $A_1, \ldots, A_n, A$, if $A_1, \ldots, A_n \vdash A$, then $A_1, \ldots, A_n \vdash_{At} A$. (p.100, [165])

This notion of weak conservativeness may be used as a necessary condition for all decent logical constants (although by the notion we cannot talk about logical constants one by one if we have more than one). However, this is not enough to make that distinction between intuitionistic logic and classical logic which is crucial for Dummett, since if we take this notion of weak conservativeness, then even classical logic formulated in the extended language is a weak conservative extension of $\vdash_{At}$.

On the other hand, the latter (stronger) notion is formulated by Garson [65] as follows. Sequent-style natural deduction systems S "are strongly conservative in the sense that the rules of S are a conservative extension of the rules written in each sublanguage of S." By using so-called valuational semantics, Garson showed that natural deduction for intuitionstic logic is strongly conservative. This result is actually favorable to Dummett. However, this is not enough, since this result depends on the particular framework of intuitionistic natural deduction.

It is well-known that if we add classical negation to a natural deduction system in intuitionistic logic, then we can derive Peirce's Law. But Peirce's Law contains only implication, so this is an example of non-conservative extension of intuitionistic logic. On this basis, some intuitionists claim that something is wrong with classical logic. However, this argument does not work if one adopts a multiple-conclusion sequent calculus, since in a multiple-conclusion one for classical logic, Peirce's Law can be derived within implicational fragment. To criticize classical logic for the lack of "global harmony", intuitionists have to provide an independent reason why multiple-conclusion sequents are problematic.[89]

Not only the notion of conservativeness does poorly serve for Dummett's aim of justifying intuitionistic logic, but it is difficult to characterize a natural class of constants that deserve to be called "logical" by means of the criterion of conservativeness, since there are some cases that are already well-entrenched as logics but show the phenomena of non-conservativeness. Here are some examples.

(1) Gödel-Dummett logic implication-disjunction-negation fragment ; adding conjunction is not conservative ([14],[83]).

(2) Adding the additive conjunction (with the standard operational rules in linear logic) to multiplicative fragment of relevant logic with mingle (RMIm) is not conservative. ([19])[90]

---

[89] Let us give some comment on a philosophical dispute on the status of these calculi concerning the issue of intuitionistic vs. classical logic. This distinction counts only in a limited sense. Tennant argued that multiple-conclusion calculi brings classicality from the back door ([165]), and Steinberger [158] objected that there are multiple-conclusion sequent calculi for intuitionistic logic, so Tennant's claim is technically problematic. This issue seems to rest on what we should mean by "multiple-conclusion sequent calculi". If we mean a sequent calculus that can have more than one formula on the succedent, then Tennant is certainly wrong, but Tennant must have been aware of the existence of sequent calculi for intuitionistic logic in this sense. The real issue rests on $R \to$ rule or R weakening. Intuitionist's critique, if any, should be more specific than a critique to the use of multiple-conclusion calculi. Also, any discussion on single- vs. multiple-conclusion sequents must take into consideration Carnap's discussion about non-standard valuation and multiple-conclusion ([31], [153].), which possibly favors multiple-conclusion sequents.

[90] RMI is a system introduced by Avron ([12],[13],[15]). RMI (partially) satisfies the

(3) Adding classical implication to intuitionistic logic ([55]).

Except the last one, which is relevant to another topic, these logics have stable positions in the hierarchy of relevant logics and superintuitionistic logics both from semantic and proof-theoretic points of view. Probably, it is worth pointing out that Gödel-Dummett logic (in the full language) has a cut-free hypersequent calculus with standard operational rules.[91] Hence, although this does not immediately show that something is wrong with the criterion of conservativeness, excluding both cases from the decent class of logics because of the non-conservativeness phenomena requires a very strong

---

variable sharing-property [17], which RM does not enjoy. Concerning subsystems of RM, we have two other cases of non-conservativeness. (1) In [1], it is reported that adding the mingle axiom $A \to (A \to A)$ to one formulation of R gives $\mathrm{RM0}_{\to}$, which is strictly weaker than another axiomatization of implication fragment of RM called $\mathrm{RM}_{\to}$. RM is not a conservative extension of $\mathrm{RM0}_{\to}$, but a conservative extension of $\mathrm{RM}_{\to}$. (2) Also, there are two implication negation fragments of RM, namely $\mathrm{RM}_{\sim,\to}$ and $\mathrm{RMI}_{\sim,\to}$. RM is a conservative extension of the former, but not a conservative extension of the latter. These cases may be more interesting than a non-conservativeness phenomenon that is well-known, namely adding additive conjunction with axioms ($(A \wedge B) \to A$, $(A \wedge B) \to B$, and $A \to (B \to (A \wedge B))$) to multiplicative relevant logic (Rm) leads to classical logic. In the last case, structural rules are hidden in the axioms. Hence, it is not too surprising that this gives classical logic. Hence, the axioms are clearly unacceptable for relevantists, but the mingle axiom is not so obviously against the relevantists' ideology. It is fair to note that Belnap may not have been interested in giving general criteria for logicality itself (his primary interest is to exclude an extreme case such as tonk). However, Belnap's argument is given under the general supposition that the meaning of the logical constants is given by specifying its inferential role. Hence, his case for excluding tonk by his criteria can readily be generalized to other cases. There is no clear reason why Belnap should have excluded these from the scope of his solution, although these phenomena were not known when [23] was written. Apparently, the realm of substructural logic is more complicated than earlier expectations.

[91] Hypersequent calculi of their fragments whose extensions to the full system are not conservative have different cut-free formulations for the full system (see, 2.5.1.). Hence, conservativeness depends on different formulations.

argument. It is extremely unlikely that Belnap would say that these conjunctions are not logical constants. These show that conservativeness as a criterion of logical constant-hood may be too strong.[92]

Conservativeness is a desirable property in most cases and non-conservativeness is usually rightly taken to be pathological, but we should not overgeneralize this. (Incidentally, in all the modal logics presented in Chapter 3, the introduction of modal operators is a conservative extension of classical propositional logic. This is guaranteed by cut-elimination.) Conservativeness can be safely used as a reason for adopting a particular formulation of logic, rather than non-conservative ones (cf. see [1]), but we have to be careful in going further.

Let us now discuss the issue of uniqueness. Like Belnap, Došen uses uniqueness as a condition of his structural analysis (2.4.2), although Došen

---

[92] One may point out that the same comment can be applied to our treatment of modal logic B or GL. They are well-entrenched modal logics in the study of modal logic. Any position that does not include these in the class of modal logics may have a very heavy burden of proof in the same way as we critically comment on conservativeness. But there is a significant difference between our case and conservativeness. In our case, we leave open a possibility that there is some way of accommodating the aforementioned cases of "modal logic" B or GL if there is some natural way of handling these cases by using some operational rules and structural rules. And also, we have never claimed that we have a positive evidence of excluding these from the class of logics. We have only stated that from our point of view we do not have sufficient reasons to consider them as logics. We have never denied that there may be some other reasons why they can be reasonably called logics. Once one adopts a mathematically precisely defined concept such as conservativeness as a criterion of logical constants, this automatically implies that the negative characterization of what is not a logical constant becomes sharp. Anything which fails to satisfy the criterion is excluded from logic. However, the boundary of logic may be intrinsically vague, so such a characterization may almost always have some problem of overgeneration or undergeneration. This may be a methodological error. Therefore, we step back to a much more modest claim about the positive test.

makes a distinction between his structural analysis and a definition (hence, their reasons for requiring uniqueness are different). Although it holds for many cases, we do not take uniqueness to be necessary for logical constanthood for the following reasons.

First, despite its name, we do not require the definitional equation to be definitions in a precise sense (traditionally understood). Belnap takes uniqueness to be necessary since he takes introducing a logical constant by rules to be assimilated to a definition. (Došen does not explicitly state the reason why uniqueness is so important.) Since we do not take it to be a definition, we do not have to require uniqueness for a logical constant.[93]

Second, Dummett states the reason why uniqueness is required as follows. "We are entitled to stipulate a set of logical laws only if we thereby fix the meaning of the logical constant that they govern"([53]). If the logical laws did not guarantee the uniqueness, then the logical laws would "not fully determine the meaning of logical constants"(ibid.) Dummett claims that uniqueness is required because otherwise the laws do not fully fix the meaning of a logical constant (i.e., specify its inferential role), and hence we are not

---

[93] Došen is among several people who explicitly claim uniqueness to be a necessary condition for a logical constant. (Wansing [173] shows uniqueness of a modal operator by using display calculi, too.) Also, Greg Restall [141] discusses (and essentially argue for) Belnap's conditions of conservativeness and uniqueness by using some cases from various logical systems including his version of hypersequent calculus for S5. It seems that Restall's claim of uniqueness of modal operator in his hypersequent calculus holds only if we give up embedding of hypersequents into the object language, since, whichever we choose, the translated rule for modality may not work unless we explicitly postulate $\vdash \Box_1 A \leftrightarrow \Box_2 A$ beforehand. But this is supposed to be *derived* in a combined system. However, it may be inevitable since in a Hilbert-style system we cannot prove uniqueness anyways.

entitled to stipulate them. We do not adopt this view, but we assume that the metalanguage is more basic (with Sambin et al. and Došen). Due to this difference, we may not have to a priori exclude a possibility that logical laws do not fully determine the meaning of a logical constant.

Third, Došen and Schroeder-Heister [46] point out that there is a kind of duality between conservativeness and uniqueness. They consider a “dual” case of tonk, which has and-introduction rule and or-elimination rule. They state that although tonk violates conservativeness but trivially satisfies uniqueness since it trivializes the system, this constant satisfies conservativeness but uniqueness does not hold. This case illustrates their notion of duality. They also state that if we require any of these conditions for a logical constant, then it makes more sense to require both of these conditions simultaneously. This makes puzzling Došen’s position of adopting only uniqueness as a criterion without requiring conservativeness (2.4.2).

Without conservativeness, uniqueness may not always be a safe criterion. Humberstone [85] uses an example (3) of adding classical implication ($\rightarrow_c$) to intuitionistic logic (with intuitionistic implication $\rightarrow_i$) to discuss the issue. This example is often mentioned as an example of non-conservativeness, since simply adding classical implication collapse the whole system into classical logic ([55]). (In this sense, this is a phenomenon similar to the case of Peirce’s law.) Namely, we can prove “uniqueness” of $\rightarrow_i$ and $\rightarrow_c$. This already shows that proving uniqueness may not be “desirable,” just because those who discuss combining logics tend to hope that different logics can be nicely combined without interfering with each (see [85] for more discussions on similar collapses.)[94] Compare this case with S4 strict implication logic plus classical implication. Although uniqueness does not hold for S4 strict implication, adding classical implication is a conservative extension of S4 strict implication logic (see the appendix). Now we have two logics where conservativeness holds but uniqueness fails (or vice versa). We seem to have no good reason to take only the former case as logic (symbols occurring in the formulas as logical constants) and not the latter, although Došen would have to accept such a consequence. Hence, if we give up conservativeness, then uniqueness should go, too.

Although it is only a positive test (hence it at most can say that currently we do not have any sufficient reason to accept a symbol as a logical constant when we meet a case that one may want to exclude from the class of logical constants), we recommend the solvability of definitional equations (using Avron's standard) as a positive criterion for being logical constants.[95] The

---

[94] The uniqueness issue arises when two symbols are governed by the same rules. So, this case may not look like a case in which the issue of uniqueness arises. However, without conservativeness, after adding a new connective, the case is simply one logical system where the two symbols follows the same rules in the sense that all the rules for $\to_c$ are derived rule for $\to_i$. Incidentally, there is a way to restore conservativeness [55]. To avoid collapsing, it suffices to restrict $R_{\to_i}$ rule to a case, where the context Γ^* is restricted to persistent formulas. Here a formula is persistent if every occurrence of classical implication is inside of the scope of intuitionistic implication. Also, let us point out in passing that although we use this example here since this is a well-known example of non-conservativeness, the philosophical significance of it must be limited since there is another very concise way of combining intuitionistic logic and classical logic [97], which satisfies conservativeness. We introduce two sorts of propositional variables and we postulate the law of excluded middle only for one sort of propositional variables. For the other one, we keep intuitionistic logic.

[95] We have recently found out that a similar point has already been made by Bonnay and Simmnauer ([26]). Apparently, they imply that solvability of the definitional equation

solvability of the definitional equations can cover almost all the commonly used propositional logical constants, including modal operators whose rules in (hyper)sequent calculi can be obtained by the method described in 2.4.3.2. Our view of logic may have the following advantages. First, we allow some open-endedness to logicality. The border of what are "logical" or "logical constants" may be intrinsically vague. Being able to formulate by some patterns of rules may fit well with a situation like this, since we may be able to formulate proof systems with desired properties for "logic" some time in the future by slightly extending a proof-theoretic framework. Second, due to the issue of cut-elimination, it seems better to have a picture that allows a distinction between logics and good logics. Then, we can say cut-elimination or even conservativeness is a feature of good logics, but we do not have to exclude some candidates of logical constants from the class of logical constants just because of non-conservativeness. Third, the characterization of logical constants is sufficiently local. The rules used for solving the equation come from the minimal property of the consequence relation.

What of harmony for intuitionists? We have no objection to the notion of invertibility or inversion principle, which is the same notion as Dummett's intrinsic harmony.[96] This notion is strongly connected with the solvability of

---

is a necessary and sufficient condition for a symbol in the object language to be a logical constant. Having actually solved a definitional equation is indeed a sufficient condition for the symbol to be a logical constant. However, we are satisfied with its being a positive test, since it is difficult to exhaust all the possible candidates of the definitional equations for less common logical constants.

[96] Some researchers use the notion of inversion principle or invertibility, e.g., [79], [81] to exclude tonk-like connectives. (Note Hodes [81] and Schroeder-Heister [144] use inversion

definitional equations (as briefly discussed in 2.4.3.2). There are some works that use this to avoid tonk. However, it has such a limited scope that it does not hold for some logical constants (multiplicative conjunction, modal operators). That is why we do not use this to avoid tonk-like expression. However, it is readily available to intuitionists. Hence, we would rather recommend Dummett to use invertibilty to avoid tonk and give up the notion of total harmony, since the role that the notion of total harmony plays in Dummett's revisionist program is unclear.[97] Perhaps, Dummett does not use invertibility because this is not enough to guarantee conservativeness. However, conservativeness may not be necessary to exclude tonk. For that particular purpose, invertibility may be sufficient for intuitionists. Dummett may have needed conservativeness to criticize classical logic. However, conservativeness works no better, for to make a clear distinction, Dummett has to give a reason

---

and uniqueness.) Also, let us mention Avron's recent solution [9] to the tonk problem. This is one of the mathematically most sophisticated solution to the problem. Avron goes down to a very fundamental level in the construction of a logical system and formulates proof systems called "canonical systems" in which most desiderata of proof systems hold including conservativeness and invertibility, and moreover we can formulate a natural condition (called "coherence") by which a connective like tonk can be excluded. The approach is interesting, but currently the scope of the work is limited to intuitionistic and classical logic. It should be extended to more general cases so that the solution can appropriately locate those non-conservative cases (RMIm, Gödel-Dummett logic) mentioned above which were discovered by Avron himself. These should be treated differently from tonk.

[97] Read [134] goes further in this direction. Read refers to Lorenzen's formulation of inversion principle, which is simply the matching of the introduction rule and the elimination rule. Following this guiding idea, Read and others introduce the idea of rules in natural deduction "generalized elimination rules." ([146], [171]). Read shows that harmony as inversion does not imply either conservative extension (even consistency) or normalization. Independent of its philosophical importance, this seems to be worth doing. In particular, this leads to a natural question of which generalized elimination rules have their sequent counterpart L-rule in sequent calculus which allows cut-elimination. (Such a Gentzenization was first done by Avron [11].)

why multiple-conclusion is wrong since classical multiple-conclusion sequent calculus enjoys the aforementioned strong conservativeness ([65]). Making a contrast between the two logics along this line may be more difficult than one might initially think. Moreover, unlike Dummett, Prawitz suggests that the notion of harmony should be understood as inversion. Discussing the notion of "canonical proofs" in the context of BHK interpretation for implication (procedure of transforming a canonical proof of the antecedent A to that of B), Prawitz states as follows.

> For instance, such a procedure may be definable in an extension of a certain language without being definable in the language itself, and hence, in this respect, the extension of a language obtained by introducing new logical constants may not be conservative extension of the original language. (p.29, [125])

It may be unclear what example Prawitz had in mind here.[98] However, Prawitz's idea about the status of conservativeness is clear enough. In fact, Prawitz has ([125], [127]) consistently been skeptical about the status of the notion of conservativeness in proof-theoretic semantics.[99] This seems to be more coherent than Dummett's adoption of two notions of harmony.

---

[98] Also, in order to precisely formulate conservativeness in the standard sense, Prawitz needs to formulate theories in these languages.

[99] Here is another criticism of Prawitz's against Dummett's view of harmony. "Dummett also suggests (pp.217-220) that the requirement of harmony between the two aspects of the use of an expression can be made precise by saying that it is equivalent to requiring that the

2.5 What is a good proof system ?

In this section let us return to more formal and methodological issues. Our discussion on Došen's principle has been mainly based on sequent calculi, but we eventually argue that hypersequent calculi are even better than traditional sequent calculi and at least not worse than other generalizations of sequent calculi. Our argument concerning modal operators rests on the use of hypersequents; hence, we have to argue for the adoption of them.

Generalizations of sequent calculi are motivated since it turns out to be difficult to formulate cut-free traditional sequent calculi for quite a few logics (including some modal logics already mentioned.) Avron says as follows.

> Powerful as it is, the framework of ordinary sequents is not capable of handling all interesting logics. There are logics with nice, simple semantics and obvious interest for which no decent, cut-free formulation seems to exist .... It would be an exaggeration to reject them as worthless just because of this fact. Larger, but still satisfactory frameworks should be sought. (p.3, [18])

---

addition of the expression to a language should not license a use of the old vocabulary which was not already licensed in the original language. This can hardly be correct, however, because from Gödel's incompleteness theorem we know that the addition to arithmetic of higher-order concepts may lead to an enriched system that is not a conservative extension of the original one in spite of the fact that some of these concepts are governed by rules that must be said to satisfy the requirement of harmony." ([128], p.374). The issue seems to rest on how they consider the continuity of logic and mathematics, but we do not go into the further details of the topic here.

Indeed, different generalizations of sequent calculi and the desiderata of a good proof theoretical framework extending sequent calculi have been proposed. In order to evaluate these proposals, it would be helpful to first discuss what are desirable features of a proof-theoretical framework (or proof systems) in general. For our purpose, Avron's and Wansing's proposals are particularly helpful as sources for our discussion. Wansing ([172]) focuses on the desirable features of operational rules.

> (1) "In the first place, the meaning assignment should not make the meaning of $\sharp$ dependent on the meaning of other connectives. That is to say, the sequent rules for $\sharp$ should be given a purely structural account of $\sharp$'s meaning, in the sense that they should not exhibit any connective other than $\sharp$." (p.128, [172] This property may be called *separation.*)
>
> (2) "Moreover, the rules for $\sharp$ should be *weakly symmetric*; every rule should either belong to a set of rules ($\sharp \Rightarrow$) which introduce $\sharp$ into premises (i.e. on the left side of $\Rightarrow$ in the conclusion sequent) or to a set of rules ($\Rightarrow \sharp$) which introduce $\sharp$ into conclusions (i.e. on the right side of $\Rightarrow$ in the conclusion sequent). The sequent rules for $\sharp$ can then be called *symmetric*, if they are weakly symmetric and both ($\Rightarrow \sharp$) and ($\sharp \Rightarrow$) are non-empty." (ibid.)
>
> (3) "The sequent rules for $\sharp$ will be called *weakly explicit*, if the

> rules ($\Rightarrow \sharp$) and ($\sharp \Rightarrow$) exhibit $\sharp$ in their conclusions only, and they will be called *explicit*, if in addition to being weakly explicit, the rules in ($\Rightarrow \sharp$) resp. ($\sharp \Rightarrow$) exhibit only one occurrence of $\sharp$ on the right resp. the left side of $\Rightarrow$." (ibid.) ($\Rightarrow$ is our notation for sequents.)

The above points are general ones for operational rules, and these points can be understood as sequent versions of the conditions originally stated for natural deductions. Wansing emphasizes the three points on operational rules because "in sequent systems whose rules are separate, symmetric, and weakly explicit the redundancy of cut implies the subformula property." (ibid.)

These are standard conditions of well-behaved operational rules, which are commonly accepted. Hence, we do not have many things to say about this except some crucial points about modal cases. Let us give just a few comments on them. Concerning (1), strictly speaking, not exhibiting any connective other than $\sharp$ does not imply that the rule is purely structural if we take the phrase "purely structural" in the same way as Došen understand in his definition (see $R \Longrightarrow$). Our (operational) modal rules all satisfy (1). Modal rules for S4 and stronger logics are symmetric, although those for weaker modal logics (K, D, T, 4) violate symmetry. More importantly, our $R\Box$ (for 4) and $R \Longrightarrow$ violate weak explicitness (3). However, we have already argued that $R \Longrightarrow$ may have some conceptual advantage, in spite of its methodological problem from a purely structural point of view.

About structural rules, Wansing [172] states as follows

> What one would need, it seems, is an extension of the usual Gentzen format that (i) conforms to the usual philosophy of meaning present in studies inspired by Gentzen and (ii) offers sufficient degrees of structural freedom. The proof theory we are in search of should exemplify a principle that has most emphatically been advocated by Došen ([45]) and that may therefore be called Došen's principle. (p.128, [172])

As a general discussion concerning structural features of logics, this may not look too much different from Došen's view of structural features of logics. However, its specific content should be more carefully examined since as discussions in 2.3 showed, there is no standard reading of Gentzen's text. Wansing has his own reading, which emphasizes schematic nature of introduction rules ([173], see the appendix). Wansing's article presents modular formulations of major normal modal logics.[100] Wansing's formulation of introduction rules for modal operators can be contentious, since to achieve this purely schematic nature, Wansing uses an approach called display calculi and

---

[100] Wansing gives the following comment on modal logics, "in contrast to the axiomatic approach, the standard sequent-style proof theory for normal modal logic fail to be 'modular', and the very mechanism behind the small range of known possible variations is not clear." (p.128, [172])

in display calculi very strong structural principles are assumed. Hence, the issue of whether this solution is satisfactory or not may require a substantial discussion on what structural rules are. We return to this issue in 2.5.2.

Next, we move on to Avron's often cited desiderata.

> (1) It should be able to handle a great diversity of logics of different types. ...
>
> (2) Because of the proof-theoretical nature and the expected generality, the framework should be independent of any particular semantics. One should not be able to guess, just from the form of the structures which are used, the intended semantics of a given proof system (recent frameworks for many-valued logics and for modal logics violate this principle - see below).
>
> (3) The structure used in the framework should be built from the formulae of the logic and should not be too complicated (for human understanding and for computer implementation). Most important - the subformula property they allow should be a *real* one. (footnote. A use of "structural connectives" that can arbitrarily be nested usually violates this priniciple. It seems to me that this is the weak point of Belnap's framework of Display Logic [24], [1] which otherwise has all the other properties.) ...
>
> (4) The rules of inference should have a small, fixed number of

> premises, and their application should have a local nature. In other words: the applicability of a rule should depend on the structure of the premises and not on the way they have been obtained.
>
> (5) Since there should be something common to all the various connectives, we call "conjunction", "disjunction", "implication" and "negation", the corresponding rules should be as standard as possible. The difference between logics should be due to some other rules, which are independent of any particular connective. Such rules are usually called "structural rules". (Footnote. This principle was put forward in [45]. Došen's paper contain a general characterization of the basic standard connectives. . . .
>
> (6) The proof systems constructed within the framework should give us better understanding of the corresponding logics and the difference between them. (pp.2 [18])

Let us first state that we basically subscribe to these methodological points. Actually, our structural-reflective view is formulated by considering these desiderata. We then have some preliminary comments. (Some discussions will be put off until 2.5.2.) Avron's point (5) is essentially the same as Došen's principle, so it does not require any additional comment (of course, except the issue of modalized context in Avron's hypersequent calculus for $\mathsf{S5}$

discussed in 2.4.4.). (4) is put here explicitly probably because there are some logics that can have a cut-free ordinary sequent calculus, but the number of premises of some rules of the calculus varies depending on the conclusion they derive (e.g., Gödel-Dummett logic [156], $\mathsf{S4.3}$ [152], GL_{lin} [170]). Also, note that there seems to be some disagreement on the way in which Avron takes the word "local" and Hacking does it, since Hacking rejects $\mathsf{S4}$ modal logic because it is not "local." Avron's notion of "local" is determined whether the applicability of a rule depends the structure of the immediate premises. In this sense, Curry-style modal rules and Avron's (and our) modal structural rules are local. But some pattern of deduction must occur above that line (see the relevant quote in 2.4.5.3.). In this sense, modal rule $R\Box$ has some non-local features. (See also the proof of cut-elimination for modal logics and our comments in Chapter 3.)

For (1), we need at least hypersequents to formulate logics because we want to cover more logics. Hypersequents can be potentially useful for both modal logics and substructural logics. However, hypersequents have been mainly used to formulate substructural logics that have no cut-free traditional sequent calculi (see 2.5.1). In modal logics, $\mathsf{S5}$ is almost the only application of hypersequents so far. In this thesis, we have extended the applications (Chapter 3 and 4).

As a proof-theoretical framework that satisfies these features, Avron recommends hypersequents, which seem to be minimal extensions of sequents. Avron states that hypersequents "form a simple and natural generalization of the sequential framework," but "it is not much more complicated and goes in fact just one step further"(p.3 [18]). In the next section, we present hypersequent calculi in a somewhat systematic manner and important technical results that can give a support for hypersequents. Then we will discuss some methodological issues.

2.5.1 Hypersequent calculi and substructural hieararchy

For the purpose of systematic exposition, we present a hypersequent calculus for Full Lambek Calculus with exchange $\mathbf{HFL_e}$ (see [34]) and its multi-conclusion variant $\mathbf{HInFL_e}$ (involutive Full Lambak calculus),[101] with basic internal and external structural rules. We also present some representative cases of characteristic external structural rules for some substructural logics as illustrations. For modal logics, we present several systems in Chapter 3. The grammar of the formulas is as follows.

$$A =:: p|0|1|\top|\bot|A_1 \to A_2|A_1 \wedge A_2|A_1 \vee A_2|A_1 \oplus A_2|A_1 \otimes A_2$$

A hypersequent is a multi-set of sequents $(\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_n \Rightarrow \Delta_n)$.

A hypersequent is interpreted precisely given by the interpretation function $(\quad)^I$.

---

[101] Although they do not use hypersequent calculi, the name comes from [64].

Definition 2.5.1 *a.* $(A_1, \ldots, A_n \Rightarrow B_1, \ldots, B_m)^I = A_1 \otimes \cdots \otimes A_n \to B_1 \oplus \cdots \oplus B_m$

b. $(\Rightarrow) = 1 \to 0$*; c.* $(S_1| \ldots |S_n)^I = ((S_1^I) \wedge 1) \vee \cdots \vee ((S_n^I) \wedge 1)$

Note: "|" is interpreted as additive disjunction and "," on a succedent is interpreted as multiplicative disjunction. For single-conclusion cases, we modify rules as follows. For $R\otimes$, let $\Pi = \Lambda = \emptyset$. For $R\wedge$, $R\vee$ and $L \to$, $\Pi = \emptyset$ and $\Lambda = \{C\}$. No rules for $\oplus$. $R2 \to$ is the single-conclusion version for $R \to$.

1) **Axioms**: $G|p \Rightarrow p$ $\quad$ $G| \Rightarrow \mathbf{1}$ $\quad$ $G|\mathbf{0} \Rightarrow$

$$G|\bot, \Gamma \Rightarrow \Pi \qquad G|\Gamma \Rightarrow \Pi, \mathbf{0} \qquad \frac{G|\Gamma \Rightarrow \Pi}{G|\mathbf{1}, \Gamma \Rightarrow \Pi} \qquad \frac{G|\Gamma \Rightarrow \Pi}{G|\Gamma \Rightarrow \Pi, \mathbf{0}}$$

2) **Operational rules**:

$$\mathbf{L}\otimes\ \frac{G|A, B, \Gamma \Rightarrow \Pi}{G|A \otimes B, \Gamma \Rightarrow \Pi} \qquad \mathbf{R}\otimes \frac{G|\Gamma \Rightarrow \Pi, A \qquad G|\Delta \Rightarrow \Lambda, B}{G|\Gamma, \Delta \Rightarrow \Pi, \Lambda, A \otimes B}$$

$$\mathbf{L}\wedge\ \frac{G|A_i, \Gamma \Rightarrow \Pi}{G|A_1 \wedge A_2, \Gamma \Rightarrow \Pi}\ (i = 1 \text{ or } 2) \qquad \mathbf{R}\wedge \frac{G|\Gamma \Rightarrow \Pi, A \qquad G|\Gamma \Rightarrow \Pi, B}{G|\Gamma \Rightarrow \Pi, A \wedge B}$$

$$\mathbf{L}\oplus \frac{G|A, \Gamma \Rightarrow \Pi \qquad G|B, \Delta \Rightarrow \Lambda}{G|A \oplus B, \Gamma, \Delta \Rightarrow \Pi, \Lambda} \qquad \mathbf{R}\oplus \frac{G|\Gamma \Rightarrow \Pi, A, B}{G|\Gamma \Rightarrow \Pi, A \oplus B}$$

$$\mathbf{L}\vee\ \frac{G|A, \Gamma \Rightarrow \Pi \qquad G|B, \Gamma \Rightarrow \Pi}{G|A \vee B, \Gamma \Rightarrow \Pi} \qquad \mathbf{R}\vee\ \frac{G|\Gamma \Rightarrow \Pi, A_i}{G|\Gamma \Rightarrow \Pi, A_1 \vee A_2}\ (i = 1 \text{ or } 2)$$

$$\mathbf{L}\rightarrow\frac{G|\Gamma\Rightarrow\Pi,A\quad G|B,\Delta\Rightarrow\Lambda}{G|A\rightarrow B,\Gamma,\Delta\Rightarrow\Pi,\Lambda}$$

$$\mathbf{R1}\rightarrow\ \frac{G|A,\Gamma\Rightarrow B}{G|\Gamma\Rightarrow A\rightarrow B,\Delta}\qquad\mathbf{R2}\rightarrow\ \frac{G|A,\Gamma\Rightarrow B}{G|\Gamma\Rightarrow A\rightarrow B}$$

3) **Basic external structural rules**:

$$\mathbf{EW}\qquad\frac{G}{G|H}\qquad\qquad\mathbf{EC}\qquad\frac{G|\Gamma\Rightarrow\Delta|\Gamma\Rightarrow\Delta}{G|\Gamma\Rightarrow\Delta}$$

4) **Internal structural rules (IW)** :

$$\mathbf{LW}\qquad\frac{G|\Gamma\Rightarrow\Delta}{G|A,\Gamma\Rightarrow\Delta}\qquad\qquad\mathbf{RW}\qquad\frac{G|\Gamma\Rightarrow\Delta}{G|\Gamma\Rightarrow\Delta,A}$$

$$\mathbf{LC}\qquad\frac{G|A,A,\Gamma\Rightarrow\Delta}{G|A,\Gamma\Rightarrow\Delta}\qquad\qquad\mathbf{RC}\qquad\frac{G|\Gamma\Rightarrow\Delta,A,A}{G|\Gamma\Rightarrow\Delta,A}$$

5) **Additional external structural rules**

$$\frac{G|\Gamma,\Delta\Rightarrow\Pi,\Lambda}{G|\Gamma\Rightarrow\Pi|\Delta\Rightarrow\Lambda}\ \text{(splitting)}\qquad\qquad\frac{G|\Gamma,\Delta\Rightarrow}{G|\Gamma\Rightarrow|\Delta\Rightarrow}\ \text{(weak splitting)}$$

$$\frac{G|\Gamma_1,\Delta_1\Rightarrow\Pi_1\qquad G|\Gamma_2,\Delta_2\Rightarrow\Pi_2}{G|\Gamma_1,\Delta_2\Rightarrow\Pi_1|\Gamma_2,\Delta_1\Rightarrow\Pi_2}\ \text{(communication)}$$

$$\frac{G_1|\Gamma,\Delta\Rightarrow A}{G|\Gamma\Rightarrow A|\Delta\Rightarrow A}\ \text{(weaker splitting)}$$

$$\frac{G_2|\Gamma \Rightarrow \Pi \qquad G|\Delta \Rightarrow \Lambda}{G_1|G_2|\Gamma, \Delta \Rightarrow \Pi, \Lambda}\ (\text{mingle})$$

$$\frac{G|\Gamma_1, \Gamma_2, \Gamma_3 \Rightarrow \Delta_1, \Delta_2, \Delta_3 \qquad G|\Gamma_1', \Gamma_2', \Gamma_3' \Rightarrow \Delta_1', \Delta_2', \Delta_3'}{G|\Gamma_1, \Gamma_1' \Rightarrow \Delta_1, \Delta_1'|\Gamma_2, \Gamma_2' \Rightarrow \Delta_2, \Delta_2'|\Gamma_3, \Gamma_3' \Rightarrow \Delta_3, \Delta_3'}\ (\text{merging for } L_3)$$

6) **Cut**

$$\frac{G_1|\Gamma_1 \Rightarrow \Delta_1, A|H_1 \qquad G_2|A, \Gamma_2 \Rightarrow \Delta_2|H_2}{G_1|G_2|\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2|H_1|H_2}$$

The correspondence between rules and logics is as follows: 1. IPC + splitting = classical logic, 2. IPC + weak splitting = logic of weak excluded middle, 3. IPC + Communication = Gödel-Dummett logic,[102] 4. mingle is used in RM,[103] 5. merging for L3 is used in Lukasiewicz 3-valued logic, 6. weak weakening is used in RM$_3$. (From 4, we omit exact correspondence since it requires tedious description.)

Not only can we have a greater variety of logics be handled by using hypersequent calculi, but we now have a technical reason why hypersequents should be taken as an important proof-theoretical framework in non-classical logics, thanks to recent work [34] and [36]. Their technical contribution are as follows.

1) Fixing operational rules (invertible in single-conclusion calculi) and modifying structural rules, they formulate a variety of logics from $\mathbf{FL_e}$ to clas-

---

[102] IPC + weaker splitting in the language $\{\rightarrow, \vee\}$ gives a logic which is strictly weaker than Gödel-Dummett logic. Adding $\wedge$ to this logic is not a conservative extension.[14], [85].

[103] The extension of the fragment $RM_{\rightarrow,\sim}$ to the full RM is conservative.

sical in hypersequent calculi, single-conclusion in [34] and multi-conclusion in [36].

2) By using the notion of polarity of the connectives introduced by Girard [68], they introduce the notion of substructural hierarchy of formulas ($\mathcal{P}_n$, $\mathcal{N}_n$) analogous to the hierarchy of first-order formulas (Σ_n and Π_n). Here is the definition. (We present only the single-conclusion case for illustration. Hence, there is no multiplicative disjunction in the definition.)

Definition 2.5.2 *For each $n \geq 0$, the sets $\mathcal{P}_n$, $\mathcal{N}_n$ of (positive and negative) formulas are defined as follows: (0) $\mathcal{P}_0 = \mathcal{N}_0 =$ the set of propositional variables.*

(P1) 1, $\bot$ and any formula $A \in \mathcal{N}_n$ belong to $\mathcal{P}_{n+1}$.

(P2) If $A, B \in \mathcal{P}_{n+1}$, then $A \vee B, A \otimes B \in \mathcal{P}_{n+1}$.

(N1) 0, $\top$ and any formula $A \in \mathcal{P}_n$ belong to $\mathcal{N}_{n+1}$.

(N2) If $A, B \in \mathcal{N}_{n+1}$, then $A \wedge B \in \mathcal{N}_{n+1}$.

(N3) If $A \in \mathcal{P}_{n+1}$ and $B \in \mathcal{N}_{n+1}$, then $A \rightarrow B \in \mathcal{N}_{n+1}$.

Proposition 2.5.3 *1. Every* $\mathbf{FL_e}$ *formula belongs to some $\mathcal{P}_n$ and $\mathcal{N}_n$.*

2. $\mathcal{P}_n \subseteq \mathcal{P}_{n+1}$, $\mathcal{N}_n \subseteq \mathcal{N}_{n+1}$ for every n.

Examples: $(p \rightarrow q) \vee (q \rightarrow p) \in \mathcal{P}_2$, and $\neg p \vee \neg\neg p \in \mathcal{P}_3$

3) They introduce an algorithm of transforming a given axiom of a substructural logic in a Hilbert-style formulation of the pertinent logic and producing a hyperstructural rule in a cut-free hypersequent calculus for the logic.

Structural rules in traditional sequent calculi can handle axioms whose substructural complexity is $\mathcal{N}_2$, and to handle axioms in $\mathcal{P}_3'$ or $\mathcal{P}_3$, we need hypersequents and external structural rules. ($'$ is introduced due to the lack of weakening in a logic. Note that $\mathcal{P}_3' \subsetneq \mathcal{P}_3$.) The upperbound of the complexity of the axioms for which their algorithm works is $\mathcal{P}_3$ (with weakening). There exist hypersequent calculi for logics that go beyond $\mathcal{P}_3$. For instance, Lukasiewicz continuum-valued logic has an axiom whose substructural complexity is $\mathcal{N}_3$. A hypersequent calculus for this logic uses a non-standard operational rule for → in Lukasiewicz continuum-valued logic.[104]

4) Cut-admissibility of the hypersequent calculi for those logics to which their algorithm is applicable are uniformly proven(applying Okada's method in [114]).

2.5.2 Extended Došen's principle and the significance of hypersequents

Due to such a systematic understanding about substructural rules, Gentzen's sequent calculi can now be regarded as special cases of hypersequent calculi, in stead of the latter being generalizations of the former. Hypersequents are by no means ad hoc extensions of sequent calculi to cope with minor exceptions, but they are required to have a general overview of substructural logics. (Above all, hypersequent calculi have been most successful in proof theory of

---

[104] The situation is similar to the case of modal logic B, although the situation in modal logic is not as robust as substructural logics, since the precise notion of hierarchy is yet to be found.

fuzzy logic [104].) Although the situation is less spectacular than substructural logics, the same statement also holds in modal logics (see Chapter 3). Hence, following [34], we introduce a terminology "hyperstructural," which stands for either ordinary structural features of sequent calculi or external structural features of hypersequent calculi. On these bases, we can restate Došen's two theses in hyperstructural terms. Concerning the first thesis [I], we suggest to replace the word "analyze" by "characterize" due to the differences between Došen's view and our own discussed in 2.4.2. What we mean by the word is Došen's analysis with the condition (1) relaxed and without the condition (3).

[I] A constant is *logical* if, and only if, it can be ultimately *characterized* in **hyperstructural** terms.

[II] Two logical systems are alternative if, and only if, they differ only in their assumptions on **hyperstructural** deductions.

This could now be identified as "extended Došen's principle," which is a very uncontentious extension of the original Došen's principle, as far as we talk about substructural logics. Still, concerning strict implication and modal logics, we have made more substantial changes on the content of the word "structural" from the structural-reflective point of view. Based on this understanding, we allow partial schema in the structural rules. Although we do not change the wording of the principles, this modified version is the extended Došen's principle we officially endorse.

Now we would like to compare hypersequents with other proof-theoretic frameworks, which may be more radically different from traditional sequent calculi. Let us first point out some limitations of hypersequent calculi.

1. It is not the case that all substructural logics can be uniformly formulated by means of hypersequent calculi. The cases whose hypersequent formulations are difficult include: relevant logics with distribution (e.g., R) ([1], [2]); fuzzy logic BL ([104]); some superintuitionistic logics (e.g., Kreisel-Putnam logic); modal logic B.

2. Unlike sequent calculi (Maehara's method), no method of extracting an interpolant from a cut-free proof has been known (except classical logic [97]).

Hypersequent calculi are relatively modest generalizations of the traditional sequent calculi, and other generalizations of sequent calculi are more expressive than hypersequent calculi. However, this does not necessarily mean that other systems are better than hypersequent calculi in every respect.[105] We have to be cautious in extending a stable framework such as Gentzen's sequent calculi. Other more expressive approaches extending sequent calculi have their own methodological drawbacks.

We raise a few issues concerning which some other approaches extending the traditional Gentzen-style sequent calculi are compared with hypersequent

---

[105] This comment probably applies to natural deductions and sequent calculi, sequent calculi and hypersequent calculi, too. But our specific purpose is to uniformly understand different logics from a proof-theoretic point of view.

calculi. This essentially amounts to giving comments on Avron's criteria (2) and (3) in 2.5.

First, we discuss the issue of independence of semantics, i.e. (2). Although proof systems that contain rich semantic information have higher expressive power, we take a strict distinction between systems containing machinery representing semantic information and purely syntactic systems to be a general methodological constraint in this thesis. This does not preclude using a proof system containing semantic information as a tool, but the reduction of such a proof system, e.g., prefixed tableaux, to purely syntactic system is significant from our point of view (see Ch.4).

Prefixed tableau systems [57], [75], [102] or any other labeled proof systems that directly encode semantic information into the proof systems themselves, e.g. Mints' indexed sequents [105] or Gabbay's labelled deduction system [63], etc., can cover a greater variety of logics (e.g. [102] covers 15 major modal logics in a modular manner). However, prefixes or labels are like remnants of semantics in proof-systems, so they violates one of Avron's methodological desiderata, i.e., (2).[106] On the other hand, hypersequent calculi have no direct connection to any particular semantics.

Concerning the coverage of logics (just as a matter of inclusion), prefixed tableau systems and hypersequent calculi are incomparable, although

---

[106] Negri's approach [107] is even more radical in incorporating semantic information. The proof system explicitly contains a predicate R corresponding to an accessibility relation. Not surprisingly, the framework can cover even more logics than prefixed tableau systems, but it carries semantic information more explicitly. Hence, this framework is even more remote from our methodology, although it also may work as a useful tool.

prefixed tableaux can cover more logics. In modal logics, cut-free prefixed tableau systems can cover all major (15) normal modal logics, but currently there are no such systems for S4.2 and S4.3 ([86] uses analytic cut for S4.3). Hypersequent calculi can handle S4.2 and S4.3 (Chapter 3) but cannot handle logic containing 5 (without 4) or B.

Second, the item (3) raises an issue of what structural rules of generalized sequent calculi are supposed to be. Recall that Wansing mentioned Došen's principle in the quote. Wansing's view is substantially different from others, since he adopts display calculi. Hypersequents and display calculi are worth comparing. Avron adds the following critical remark on two other approaches in a footnote. "Two frameworks which were proposed and deserve mentioning here are that of higher-order sequents [45] and Display Logic [24], [2]. Both are somewhat weak, I believe, with respect to point (3) above" (ibid.[18]). Display calculi were invented by Belnap [24]. The idea is to generalize the use of auxiliary symbols used in Gentzen's sequent calculus, namely ",", which behaves conjunctively on the antecedent and disjunctively on the succedent. Belnap introduced several different symbols that correspond to structural inferences, called "structural connectives." (In the appendix, Wansing's display calculi for modal logics are presented.) Display calculi have a great flexibility and are more expressive than hypersequent calculi.[107] Also, about the syntax/semantics distinction, it is quite obvious that display calculi are

---

[107] Wansing [174] proved some embedding results from hypersequent calculi to display logics, but there are reasons to think that embedding the other way is hopeless.

purely syntactic systems.

In spite of this advantage of display calculi over hypersequent calculi, the appraisal of display calculi is not simple. We claim that these two approaches are also incomparable. A potential problem for display calculi can be raised by the following technical result in display calculi in modal logic. About normal modal logics, we do have modular cut-free formulations in display calculi [172]. Kracht showed that for any elementary subframe logics can be formulated by cut-free display calculi. However, Kracht also showed the following result concerning the display calculi for these modal logics: it is undecidable whether or not a display calculus is decidable. In spite of their uniform formulations of these modal logics, the decidability is not implied by cut-elimination and a uniform decision procedure is out of reach. (Cut-elimination does not imply decidability, although modal logics formulated by display calculi are already known to be decidable.)

This is related to the problem that display calculi do not have the substructure property (structural symbols introduced as a generalization of punctuation may disappear in a cut-free derivation), as Avron [18] pointed out. Hence, cut-elimination does not have the same significance as it has in the traditional sequent calculi. The following difference is important. Strictly speaking, commas in sequent calculi disappear even in a cut-free proof, and the same comment applies to hypersequent calculi concerning their "|". These disappear even in a cut-free proof; however, in the case of display logic, the same phenomenon happens much more radically since "a use of structural symbols can be arbitrarily nested" ([18]). Wansing does not explicitly discuss what kind of general features structural rules in a proof system have to satisfy. (But it is fair to note that except Došen's definition we cited, there seems to be no general definition of structural rules, probably since Gentzen.)

We try to figure out why introducing structural symbols may be problematic, although this is a little speculative. To explain the situation, we introduce a new term "grades of reflection" of the (deductive) metalanguage.

0th grade of reflection : there is no specification in the object language at all (abstract consequence relations).

1st grade : purely structural rules and purely schematic operational rules are allowed. (Multiplicative and additive conjunctions and disjunctions are characterized.) [1 1/2 grade : one could separate implication here, since formulating and solving the definitional equation for implication needed some compromising.]

2nd grade : structural rules and partially schematic operational rules are allowed. Contexts are modalized. ($\mathsf{S4}$ $\Box$ and $\mathsf{S4}$ $\to$.)

3rd grade : structural rules and partially schematic operational rules are allowed. Modalized contexts can play a distinguished structural role in a system. (Exponentials ! from linear logic, modal hypersequents for $\mathsf{S5}$, $\mathsf{S4.3}$)

4th grade : nested structural connectives and purely schematic operational rules are allowed. (Display calculi for normal modal logics.) This "externalizes" modal operator $\Box$ via modal structural connective $\bullet$ into the metalanguage, where roughly $\bullet$ works for $\Box$, as "," does for conjunction $\wedge$.

(see the appendix for more details.)

There is a significant difference between 3rd and 4th grades of reflective involvement. If we take into consideration arbitrarily nested structural symbols, among constants introduced in the object language introduced via reflection of metalinguistic operations, there seems to be a natural borderline between what we can clearly call logical constants through structural characterization in the sense of extended Došen's principle and what is beyond that. We have done up to 3rd grade. As we discussed, at this level we already did not find an entirely satisfactory intuitive justification for using modal operators as structural markers. In addition to technical results, some intuitive justification for the use of $\bullet$ in display calculi is desirable. From the reflective point of view, $\bullet$ must have its origin in $\vdash$. However, it appears that $\bullet$ can do more than the original $\vdash$, since we usually do not allow an arbitrary nesting $\vdash$ with other auxiliary meta-linguistic symbol such as "$,$". Both Došen and Sambin et al. assume that the metalanguage is more basic. But in display calculi, it does not seem that the metalinguistic notion $\bullet$ is more basic than the object language $\Box$. (Note the difference between $\bullet$ and $\Box$. The latter never disappears in a cut-free proof.) Then it is probably legitimate to ask on what basis we are entitled to give such a structural capacity to structural connectives in display calculi. This does not mean that display calculi are defective, but this feature may need more scrutiny. The following remark of Sambin et al. can be understood from such a point of view.

> Our *and* and *yields* correspond to what in display logic is denoted with punctuation signs like ",", ";", "$*$" etc. and called "structural connectives." However now it should also be clear that adopting such terminology, rather than "metalinguistic links" as we did, would jeopardize all the effort to clarify the role of metalanguage. (p.990, [143])

Our critical comment can be turned into a constructive research program to find out something stable in the middle. In fact, concerning Avron's comment cited above and the issue of decidability in display calculi, Goré suggests as follows.

> Having tried to obtain decision procedures from display calculi, I must confess that I am also sometimes frustrated by this aspect of display calculi. I think that the truth probably lies somewhere in the middle. Certainly it is true that more traditional calculi give decision procedures for simple logics via the "real" subformula property. However, this usually only works for the very simplest of logics. As soon as contraction or weakening aspects surface, the problems become difficult even with a "real" subformula property as witnessed by traditional Gentzen calculi for propositional Linear Logic. In display calculi, these problems surface even at the

> base level (p.270, [76]).

We share this view with Goré. Hypersequents can be extended to a more general framework and display calculi may need some computationally well-behaved subsystems.[108]

Before moving on to the conclusion, let us add one point concerning another significance of hypersequents. Not only are hyperesequents good in the general sense discussed above, but they also turn out to be useful tools for at least some special cases of combining logics including the technical material in this thesis. Although there exist some works of this direction of research (such as [103]), the role that hypersequents might take in combining logics in a proof system has been mostly neglected. In this thesis, this aspect of hypersequents is highlighted. Unfortunately, the ability of hypersequents to

---

[108] Display calculi were introduced as generalizations of so-called Dunn-Mints calculi [118], which were introduced in order to formulate relevant logic with distribution such as R. The calculi have two punctuation marks "," and ";" that correspond to extensional and intensional contexts, respectively. This approach is taken over by Slaney [154], Read [132], and Schroeder-Heister [147], Pym and O'hearn [111] and now called "bunched calculi." However, apparently, this approach is incomparable with hypersequent calculi. The areas of applications hardly overlap.

There is yet another approach called nested sequents ([29]. See also [123]). This allows an arbitrary nesting of $\Rightarrow$ in sequent calculi. Although the authors claim that their approach is purely syntactic (based on the existence of syntactic translation from nested sequents to formulas in the object language), the point may be moot according to Avron's condition (2). Also, from a methodological point of view, it is crucial that we are hesitant to call nested sequents for different normal modal logics "structural variants," since the forms of rules look too far from structural variants. (This point may raise some disagreement and we may be brought back to the issue of what are "structural rules.") However, this approach may be interesting for the following two reasons. 1. This approach is likely to come between hypersequents and display calculi (probably via embeddings). 2. this may have an independent motivation mentioned in a footnote in the section 2.4.3.2.b.

combine logics is limited, but whenever they work, they can be a particularly good tool for combining logics.[109]

2.6 Conclusion and future directions

Having surveyed philosophy of logic motivated by Gentzen's systems, we have found Došen's structural view most important since the view takes the role of structural rules most seriously. Also, the view regarding logical constants as results of reflecting metatheoretic operations into the object language, which we call "reflective view," looks like a very natural view of logical constants. (According to this view, cut-elimination is not a necessary condition for being logic, but it is a necessary condition for being a good logic.) By combining these views and by slightly extending these views ("structural-reflective view") we have accommodated a variety of modal logics from a proof-theoretic point of view only using minimal generalizations of sequent calculi. A variety of modal logics can be taken to be hyperstructural variants, which have modal operators that are obtained by reflecting $\vdash$ of the underlying consequence relations and whose operational rules are fixed throughout these modal logics. Hence, according to this view, these modal operators are decent logical constants. As particular applications, we gave cut-free hypersequent calculi for several modal logics (in Chapter 3). We did not originally intend to consider further implications of our view, but with hindsight we

---

[109] Although it looks like a challenging project, for several reasons, subset space logic [39] looks like an interesting case to accommodate by using hypersequents.

have realized that our view also does the following : 1. introducing modal operators for a variety of modal logics without using possible worlds; 2. providing a defence of the use of Curry-style modal rules in the technical part of our thesis (Ch.4).[110] From this broader perspective, we argued that conservativeness and uniqueness should play more limited roles in philosophical discussions of foundations of logic. Lastly, we gave a defense of hypersequent calculi in comparison with other generalized sequent calculi.

The discussions of this chapter raised many further technical and conceptual problems. First of all, in order to push the reflective view of logic one step further, it makes sense to explore a construction like a nested $\vdash$ more systematically. We can also explore substructural systems for strict implications of $\mathsf{S4}$ (this will be systems around relevant logic E [1]) and consider the possibility of defining a modal or strict implication hierarchy analogous to the substructural hierarchy. $\mathsf{S4}$ strict implication can also be explored from the viewpoint of proof-theoretic semantics. We can compare $\mathsf{S4}$ strict implication as a subintuitionistic logic and intuitionistic logic from the viewpoint of Dummett's notion of (intrinsic) harmony. Curry-Howard isomorphism in this direction may also be worth studying.

---

[110] In fact, the same point applies to intuitionistic logic with classical atoms [97], since in it one external structural rule is restricted by a feature of the object language, although the two-sorted language there is probably less contentious than modal case. Also, since we have modal embedding of the logic into an intuitionistic modal logic, our use of the structural rule applied to restricted items in the object language is justified if modal cases are justified.

Chapter 3

Hypersequents for strict implications and modal logics

3.1 Hypersequents for strict implication and modal logics

In Chapter 2, we discussed proof-theoretic foundations of logics and argue that there is a sufficient reason to think that modal operators (of many normal modal logics), which can be definable from strict implications (or obtained by modifying the rules for modal operators so obtained), are logical constants. The important technical basis of our argument is that for these cases of modal logics, we have cut-free sequent calculi or cut-free hypersequent calculi. (In order to merely argue that these modal operators are logical constants, this is actually more than sufficient, since from our point of view having a cut-free sequent calculus or hypersequent calculus is actually a necessary condition of not a “logic” but a “good” logic. However, naturally, we would also like to argue that the modal logics that we can formulate in a manner described above are good logics from a proof-theoretic point of view. In this sense, cut-elimination has crucial importance.)

In order to substantiate our philosophical claims, we give some technical details of our proof-theoretic results on strict implication logics and modal logics. Here we present hypersequent calculi for some strict implication logics and modal logics. We first present Hilbert-style systems and then present hypersequent calculi.

3.1.1 Strict implication logics

We present some Hilbert-style axiomatic systems for strict implication logics. Let us first give the grammar of the language of strict implication logics. Here we use $\longrightarrow$ for strict implication for a stylistic reason (note that this is a notational difference from the main body of Chapter 2). To formulate some logics, it is convenient to add $\top$ into the language.

$$A := P_i|\bot|\top|A_1 \longrightarrow A_2|A_1 \wedge A_2|A_1 \vee A_2$$

3.1.1.1 Hilbert-style axiomatizations for strict implication logics

The main items here are normal strict implication logics including S4 strict implication. In Hilbert-style systems, the implicational fragment of SIS4

($\mathsf{SIS4}_{\longrightarrow}$) can be axiomatized by the following simple axioms and modus ponens ([1]).

$\mathsf{I1}$. $A \longrightarrow A$

$\mathsf{I2}$. $(A \longrightarrow B) \longrightarrow (C \longrightarrow (A \longrightarrow B))$

$\mathsf{I3}$. $(A \longrightarrow (B \longrightarrow C)) \longrightarrow ((A \longrightarrow B) \longrightarrow (A \longrightarrow C))$

However, we have $\wedge, \vee, \bot, \top$ in the language ($\top$ is our addition). It does not essentially extend the system, but it is convenient when we define modal operator from the strict implication (also, note that $\wedge$, $\vee$, $\bot$, $\top$ are truth functional in classical sense, but they are not functionally complete due to Post's result and the system does not contain classical logic). We need axioms and rules as follows (in [37] except $\top$). The following system that consists of 1-10 and R1-R3 is the most basic system which axiomatizes strict implication logics for K, which is formulated by Corsi. To be precise, it is a slight modification of Corsi's systems, R1 and R3 are combined as a single rule, but by R1 and R2 it is derivable.

1. $A \longrightarrow A$
2. $((A \longrightarrow B) \wedge (B \longrightarrow C)) \longrightarrow (A \longrightarrow C)$
3. $((A \longrightarrow B) \wedge (A \longrightarrow C) \longrightarrow (A \longrightarrow (B \wedge C))$
4. $(A \wedge B) \longrightarrow A$; 5. $(A \wedge B) \longrightarrow B$
6. $A \longrightarrow (A \vee B)$; 7. $B \longrightarrow (A \vee B)$

8. $((A \longrightarrow C) \wedge (B \longrightarrow C)) \longrightarrow ((A \vee B) \longrightarrow C)$

9. $(A \wedge (B \vee C)) \longrightarrow (A \wedge B) \vee (A \wedge C)$

10. $\bot \longrightarrow A.$

11. $A \longrightarrow \top.$

R1. $\dfrac{\vdash A \qquad \vdash A \longrightarrow B}{\vdash B}$ R2. $\dfrac{\vdash A}{\vdash B \longrightarrow A}$ R3. $\dfrac{\vdash A \qquad \vdash B}{\vdash A \wedge B}$

For reflexivity, we need the following additional axiom.

IT. $A \wedge (A \longrightarrow B) \longrightarrow B$

We take the 1 - 11 plus I2 and I4 to be an axiomatization of SIS4. SIS4.3, and SIS5 can be axiomatized by adding the following axioms to SIS4, respectively.[1]

IS4.3: $((A \longrightarrow B) \longrightarrow (C \longrightarrow D)) \vee ((C \longrightarrow D) \longrightarrow (A \longrightarrow B))$

I5: $((A \longrightarrow B) \longrightarrow C) \vee (A \longrightarrow B)$.[2]

Remarks:

(a) Actually, I5 exactly corresponds to the Euclidean property of a frame.

---

[1] Corsi's formulation of the axiom for connectedness is slightly different from the one given here. Corsi's axiom can be derived from our axiom S4.3. We leave SIS4.2 for future investigations.

[2] Equivalently, $((A \longrightarrow B) \longrightarrow \bot) \vee (A \longrightarrow B)$.

Incidentally, Corsi [37] gives a few more axioms for strict implication. E.g.,

$A \longrightarrow (B \vee ((A \longrightarrow B) \longrightarrow \bot))$ (symmetry)[3]

(b) We can formulate a principle corresponding to truth persistence in the object language by using $A \longrightarrow (B \longrightarrow A)$.[4] Adding this to implicational fragment of SIS4 (I1-I3) is the implicational fragment of intuitionistic propositional logic ([1]).

In a similar way, by adding $A \longrightarrow (B \longrightarrow A)$ (TP: truth persistence) to the axiom systems of SIS4.3, SI5, we can prove the following formulas as theorems.

$\mathsf{SIS4.3} + \mathrm{TP} \vdash (A \longrightarrow B) \vee (B \longrightarrow A)$
$\mathsf{SIS5} + \mathrm{TP} \vdash (A \longrightarrow \bot) \vee A$

Obviously, these are the principle of Gödel-Dummett axiom and the law of excluded middle. This means that these strict implication logics can be taken as sub-Gödel-Dummett logic and sub-classical logic, respectively, in the sense of sub-intuitionistic logic introduced in [137]. Our presentation in this direction is not very systematic at this point, but this is one of possible directions of future research that have come out of our proof-theoretic investigations of modality.

---

[3] Equivalently, $A \longrightarrow (((A \longrightarrow B) \longrightarrow C) \vee B)$ (symmetry)

[4] An axiom for truth persistence in the sequent format is given in [87]: $A \Rightarrow B \longrightarrow A$.

We leave out semantic completeness of the Hilbert-style systems for strict implication logics here, since all the systems are covered by [37] or (in a little different style) [87].

3.1.1.2 Hypersequent calculi for strict implication logics

We now present hypersequent calculi for strict implication logics including S4 (SIS4, SIS4.3, SIS5). We confine our presentation here to simple cases.

1) Operational rules (or an axiom) for classical conjunction, disjunction, and falsum are fixed (in hypersequents). 2) Basic internal and external structural rules are fixed. These items can be presented as follows.

1) **Axioms**:[5] $p \Rightarrow p$ $\quad \bot \Rightarrow$ $\quad \Rightarrow \top$

2)**Operational rules**:

$$\mathbf{L}\wedge \ \frac{G|A, \Gamma \Rightarrow \Pi}{G|A \wedge B, \Gamma \Rightarrow \Pi} \qquad \frac{G|B, \Gamma \Rightarrow \Pi}{G|A \wedge B, \Gamma \Rightarrow \Pi}$$

$$\mathbf{R}\wedge \frac{G|\Gamma \Rightarrow \Pi, A \qquad G|\Gamma \Rightarrow \Pi, B}{G|\Gamma \Rightarrow \Pi, A \wedge B}$$

$$\mathbf{L}\vee \ \frac{G|A, \Gamma \Rightarrow \Pi \qquad G|B, \Gamma \Rightarrow \Pi}{G|A \vee B, \Gamma \Rightarrow \Pi} \qquad \mathbf{R}\vee \ \frac{G|\Gamma \Rightarrow \Pi, A_i}{G|\Gamma \Rightarrow \Pi, A_1 \vee A_2} \ (i = 1 \text{ or } 2)$$

---

[5] For some technical reason, we use axioms of the restricted form that have only propositional variables. Axioms of the general form can be derived.

3) **Basic external structural rules**

EW $$\frac{G}{G|H}$$ **EC** $$\frac{G|\Gamma \Rightarrow \Delta|\Gamma \Rightarrow \Delta}{G|\Gamma \Rightarrow \Delta}$$

4) **Internal structural rules (IW)** :

LW $$\frac{G|\Gamma \Rightarrow \Delta}{G|A, \Gamma \Rightarrow \Delta}$$ **RW** $$\frac{G|\Gamma \Rightarrow \Delta}{G|\Gamma \Rightarrow \Delta, A}$$

Cut $$\frac{G|\Gamma_1 \Rightarrow \Delta_1, A \qquad G|A, \Gamma_2 \Rightarrow \Delta_2}{G|\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2}$$

Note: 1. This is the basic system common to all hypersequents presented in this chapter.

2. We assume that a sequent is a pair of sets of formulas, and a hypersequent is a multiset of sequents. Hence, internal contractions (LC, RC) are absorbed in this assumption.

5) **Rules for strict implication logics**

SIS4 cases are the basic ones common to other cases presented below. Here we use the notation $\Gamma_{\rightarrow}$, which means that this is the set of formulas whose outermost logical symbols are $\longrightarrow$.

SIS4 $$\frac{G|\Gamma_{\rightarrow}, A \Rightarrow B}{G|\Gamma_{\rightarrow} \Rightarrow A \longrightarrow B} \qquad \frac{G|\Gamma \Rightarrow \Delta, A \qquad G|B, \Gamma \Rightarrow \Delta}{G|A \longrightarrow B, \Gamma \Rightarrow \Delta}$$

In hypersequent calculi for other SI logics, the following "external structural rules" are added to SIS4.

1. SIS4.3 (si-comm): $\dfrac{G|\Gamma, \Theta_{\rightarrow} \Rightarrow \Delta \qquad G|\Sigma, \Pi_{\rightarrow} \Rightarrow \Lambda}{G|\Gamma, \Pi_{\rightarrow} \Rightarrow \Delta|\Sigma, \Theta_{\rightarrow} \Rightarrow \Lambda}$

2. SIS5 (si-splitting): $\dfrac{G|\Gamma_{\rightarrow}, \Pi \Rightarrow \Sigma}{G|\Gamma_{\rightarrow} \Rightarrow |\Pi \Rightarrow \Sigma}$

Note : Sequent calculi for strict implication logics (strictly) weaker than S4 are quite complicated. For instance, the rule for K4 looks as follows ([87]).

$$\text{K4}\longrightarrow \text{ rule: } \frac{\Pi, \Gamma_1, A \Rightarrow B, \Delta_1 \quad \ldots \Pi, \Gamma_{2^n}, A \Rightarrow B, \Delta_{2^n}}{C_1 \longrightarrow D_1, \ldots, C_n \longrightarrow D_n \Rightarrow A \longrightarrow B} \qquad (n \geq 0)$$

where 1. $\Pi = \{C_1 \longrightarrow D_1, \ldots, C_n \longrightarrow D_n\}$. 2. $\Gamma_i = \{C_j | j \in \gamma(i)\}$ and $\Delta_i = \{D_j | j \in \delta(i)\}$. $\gamma(i)$ and $\delta(i)$ are sets of natural numbers defined as follows. Enumerate all the subsets of $\{1, \ldots, n\}$ in ascending order according to the size. Let $\gamma(i) =$ ith subset and $\delta(i) = \{1, \ldots, n\} \setminus \gamma(i)$.

Although this may not be the only way to formulate weaker strict implication logics, this is at least one evidence that S4 strict implication (and stronger ones) are special cases where rules are so simple that the connection to modality can be easily made.

3.1.2 Modal logics

Here again we first present Hilbert-style axiomatic systems and then give hypersequent (sequent) calculi. The common language of all the modal logics we discuss in this paper is specified as follows. (In modal logic, we use the the symbol → to stand for material implication. Although this may look confusing, confusion is practically unlikely due to the contexts of use. Also, we add "¬" as the primitive symbol in the language to formulate some axioms.)

$$A := P_i|\bot|\neg A|A_1 \to A_2|A_1 \wedge A_2|A_1 \vee A_2|\Box A$$

3.1.2.1 Hilbert-style axiomatic systems for modal logics

We present Hilbert-style axiomatic systems for the modal logics that can also be formulated via the basic modal rules for $\mathsf{S4}$ (or $\mathsf{K4}$, etc.) and characteristic external structural rules in hypersequent calculi. Our hypersequent formulations of these logics are given as applications of philosophical ideas on modal logics in Chapter 2.

Axioms

0) Axioms of Propositional Logic

1) Axioms of $\mathsf{K4}$: K axiom $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$ and $\mathsf{4}$ axiom $\Box\varphi \to \Box\Box\varphi$

2) Axioms of $\mathsf{KD4}$: K axiom $\mathsf{K4}$ axioms + D axiom: $\neg\Box\bot$ (or $\Box\varphi \to$

$\neg\Box\neg\varphi)$

3) Axioms for $\mathsf{wS4}$: $\mathsf{K4}$ + weak density $\Box\Box\varphi \to \Box\varphi$

4) Axioms for $\mathsf{S4}$: $\mathsf{K4}$ + reflexivity (T) : $\Box\varphi \to \varphi$

5) Axioms for $\mathsf{S4.2}$: $\mathsf{S4}$ + $.2$ (over $\mathsf{S4}$ or G^{+})[6]: $\neg\Box\neg\Box\varphi \to \Box\neg\Box\neg\varphi$

6) Axioms for $\mathsf{S4.3}$: $\mathsf{S4}$ + $.3$ (over $\mathsf{S4}$ or L^{+}) : $\Box(\Box\varphi \to \psi) \vee \Box(\Box\psi \to \varphi)$

7) Axioms for $\mathsf{S5}$: $\mathsf{S4}$ + $\mathsf{5}$: $\neg\Box\varphi \to \Box\neg\Box\varphi$

8) Axioms for $\mathsf{KD45}$: $\mathsf{KD4}$ + $\mathsf{5}$: $\neg\Box\varphi \to \Box\neg\Box\varphi$

Rules of Inference: Modus Ponens $\dfrac{A \to B \quad A}{B}$ Necessitation $\dfrac{\varphi}{\Box\varphi}$

Note: 1. The axiom called $.2$ over $\mathsf{S4}$ (G^{+}) corresponds to the frame condition "strongly directed," and the axiom called $.2$ over K corresponds to the frame condition corresponds to "directed."

2. L^{+} corresponds to the frame condition "strong connectedness."

3.1.2.2 Hilbert-style axiomatizations of combined modal logics

We can extend the language of propositional modal logic by another modal operator $\boxdot$ and formulate bimodal logics that combine the logics we discussed above. In particular, we can consider combinations of proof-theoretically well-behaved modal logics that include $\mathsf{S4}$ ($\mathsf{S4}$, $\mathsf{S4.2}$, $\mathsf{S4.3}$, $\mathsf{S5}$) by using a very simple combining axiom. We can consider six combinations of these logics. However, it turns out that only $\mathsf{S4} + \mathsf{L}$, where $\mathsf{L} = \mathsf{S4.2}$, $\mathsf{S4.3}$, or $\mathsf{S5}$,

---

[6] $\Diamond\Box\varphi \to \Box\Diamond\varphi$ if we use $\Diamond$ in the language, but we do not consider $\Diamond$ in this paper.

can have simple formulations in hypersequent calculi. The grammar of the language looks as follows.

$$A := P_i|\bot|\neg A|A_1 \to A_2|A_1 \wedge A_2|A_1 \vee A_2|\Box A| \boxdot A$$

We keep axioms and rules from each logic. The only connecting axiom added to the three combined modal logics in Hilbert-style axiomatization is as follows([56]):

$$\boxdot A \to \Box A$$

3.1.3 Sequent and Hypersequent calculi for modal logics

3.1.3.1 Basic modal rules

The standard structural and operational rules (for $\wedge$, $\vee$, $\bot$) are the same as the strict implication logics. We add the following rules for material implication. For the sake of simplicity, in sequent calculi or hypersequent calculi for modal logics, we identify $\neg A$ with $A \to \bot$. We do not take $\neg$ to be primitive.

$$L\to \ \frac{G|\Gamma \Rightarrow \Delta, A \qquad B, \Gamma \Rightarrow \Delta}{G|A \to B, \Gamma \Rightarrow \Delta} \qquad\qquad R\to \ \frac{G|\Gamma, A \Rightarrow \Delta, B}{G|\Gamma \Rightarrow \Delta, A \to B}$$

The modal rules for $\mathsf{S4}$ are the most basic ones. Others can be obtained by slightly weakening the rules via manipulating contexts.[7]

Let us first present cut-free *sequent* calculi that can be obtainable based on the idea of reflection presented in Chapter 2 for the completeness of presentation (we do not prove theorems for these here).

Rule for K4[8] $\dfrac{\Box\Gamma, \Delta \Rightarrow \varphi}{\Box\Gamma, \Box\Delta \Rightarrow \Box\varphi}$ Rule for KD4[9] $\mathrm{L}\Box_{D4}\ \dfrac{\Gamma, \Box\Delta \Rightarrow}{\Box\Gamma, \Box\Delta \Rightarrow}$

Rule for K45, KD45 $\dfrac{\Box\Gamma, \Theta \Rightarrow \Pi, \Box\Delta}{\Box\Gamma, \Box\Theta \Rightarrow \Box\Pi, \Box\Delta}$

Here Π has at most one formula in it, and KD45 is obtained by allowing Π, $\Box\Delta$ to be empty.

Rules for S4 $L\Box\ \dfrac{G|A, \Gamma \Rightarrow \Delta}{G|\Box A, \Gamma \Rightarrow \Delta}$ $R\Box\ \dfrac{G|\Box\Gamma \Rightarrow \varphi}{G|\Box\Gamma \Rightarrow \Box\varphi}$

---

[7] We can also have weak density rule by using a similar method: Weak Density: $\dfrac{\Gamma, \Box\Delta \Rightarrow \Box\varphi}{\Box\Gamma, \Box\Delta \Rightarrow \Box\varphi}$.
A proof will not be too complicated, but we state cut-elimination of this case as a conjecture.

[8] The K rule is obtained by making $\Box\Gamma$ empty.

[9] The D rule is obtained by making Γ empty.

3.1.3.2 Modal external structural rules

Restricted Modal Splitting

$$\dfrac{G|\Box\Gamma, \Box\Delta \Rightarrow}{G|\Box\Gamma \Rightarrow |\Box\Delta \Rightarrow}$$

Modal Communication

$$\dfrac{G|\Sigma, \Box\Gamma \Rightarrow \Pi \quad G|\Theta, \Box\Delta \Rightarrow \Lambda}{G|\Sigma, \Box\Delta \Rightarrow \Pi|\Theta, \Box\Gamma \Rightarrow \Lambda}$$

Modal Splitting

$$\dfrac{G|\Box\Gamma, \Delta \Rightarrow \Pi}{G|\Box\Gamma \Rightarrow |\Delta \Rightarrow \Pi}$$

Example: S4.3$\vdash \Box(\Box(\Box A \vee \Box B)) \to A) \vee \Box(\Box(\Box A \vee \Box B) \to B)$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A \Rightarrow A}{\Box A \Rightarrow A}}{\Box A \Rightarrow A|\Box A \Rightarrow B} \quad \dfrac{\dfrac{A \Rightarrow A}{\Box A \Rightarrow A} \quad \dfrac{B \Rightarrow B}{\Box B \Rightarrow B}}{\Box B \Rightarrow A|\Box A \Rightarrow B}}{\Box A \vee \Box B \Rightarrow A|\Box A \Rightarrow B} \quad \dfrac{\dfrac{B \Rightarrow B}{\Box B \Rightarrow B}}{\Box A \vee \Box B \Rightarrow A|\Box B \Rightarrow B}}{\Box A \vee \Box B \Rightarrow A|\Box A \vee \Box B \Rightarrow B}}{\Box(\Box A \vee \Box B) \Rightarrow A|\Box(\Box A \vee \Box B) \Rightarrow B}$$

3.1.3.3 Rules for combined logics

We present combined logics $\mathsf{S4} + \mathsf{L}$, where $\mathsf{L} = \{\mathsf{S4.2}, \mathsf{S4.3}, \mathsf{S5}\}$. (Due to the issue of how we should interpret hypersequents, it turns out that $\mathsf{S4.2}$ + $\mathsf{S4.2}$, $\mathsf{S4.3}$ + $\mathsf{S4.3}$, $\mathsf{S5}$ + $\mathsf{S5}$ are difficult to formulate by hypersequents.[10] $\mathsf{S4}$ + $\mathsf{S5}$ can be syntactically taken as a hypersequent formulation of Goranko

---

[10] These combinations have a problem even if we use $\boxdot$ for a translation into the object language. In some step of proving soundness w.r.t. the Hilbert-style system for $\mathsf{S4.2}$ + $\mathsf{S5}$, for instance, apparently requires $\Box A \to \boxdot\Box A$, but this is not provable in the system, which can be checked by an easy model-theoretic argument.

and Passy's logic with universal modality ([74]).

$$L\Box\ \frac{G|A,\Gamma \Rightarrow \Delta}{G|\Box A,\Gamma \Rightarrow \Delta} \qquad L\boxdot\ \frac{G|A,\Gamma \Rightarrow \Delta}{G|\boxdot A,\Gamma \Rightarrow \Delta}$$

$$R\Box\ \frac{G|\Box\Gamma,\boxdot\Delta \Rightarrow \varphi}{G|\Box\Gamma,\boxdot\Delta \Rightarrow \Box\varphi} \qquad R\boxdot\ \frac{G|\boxdot\Gamma \Rightarrow \varphi}{G|\boxdot\Gamma \Rightarrow \boxdot\varphi}$$

$$\mathsf{S4.2}\ \frac{G|\boxdot\Gamma,\boxdot\Delta \Rightarrow}{G|\boxdot\Gamma \Rightarrow |\boxdot\Delta \Rightarrow} \qquad \mathsf{S4.3}\ \frac{G|\Theta,\boxdot\Gamma \Rightarrow \Pi \quad G|\Xi,\boxdot\Delta \Rightarrow \Lambda}{G|\Theta,\boxdot\Delta \Rightarrow \Pi|\Xi,\boxdot\Gamma \Rightarrow \Lambda}$$

$$\mathsf{S5}\ \frac{G|\boxdot\Gamma,\Delta \Rightarrow \Pi}{G|\boxdot\Gamma \Rightarrow |\Delta \Rightarrow \Pi}$$

Example : $\mathsf{S4}+\mathsf{S5} \vdash \Box(\boxdot(\Box A \vee \Box B) \to A) \vee \Box(\boxdot(\Box A \vee \Box B) \to B)$.

Let us put this formula as Φ in the following derivation.

$$\begin{array}{c}
A \Rightarrow A \\ \hline
\Box A \Rightarrow A \\ \hline
\boxdot(\Box A \vee \Box B), \Box A \Rightarrow A \\ \hline
\Box A \Rightarrow \boxdot(\Box A \vee \Box B) \to A \\ \hline
\Box A \Rightarrow \Box(\boxdot(\Box A \vee \Box B) \to A) \\ \hline
\Box A \Rightarrow \Box(\boxdot(\Box A \vee \Box B) \to A) \vee \Box(\boxdot(\Box A \vee \Box B) \to B) \\ \hline
(1) \Box A \Rightarrow \Phi
\end{array}$$

$$\begin{array}{c}
B \Rightarrow B \\ \hline
\Box B \Rightarrow B \\ \hline
\boxdot(\Box A \vee \Box B), \Box B \Rightarrow B \\ \hline
\Box B \Rightarrow \boxdot(\Box A \vee \Box B) \to B \\ \hline
\Box B \Rightarrow \Box(\boxdot(\Box A \vee \Box B) \to B) \\ \hline
\Box B \Rightarrow \Box(\boxdot(\Box A \vee \Box B) \to A) \vee \Box(\boxdot(\Box A \vee \Box B) \to B) \\ \hline
(2) \Box B \Rightarrow \Phi
\end{array}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Rightarrow (1)\Box A \Rightarrow \Phi \qquad \Rightarrow (2)\Box B \Rightarrow \Phi}{\Box A \vee \Box B \Rightarrow \Phi}}{\boxdot(\Box A \vee \Box B) \Rightarrow \Phi}}{\boxdot(\Box A \vee \Box B) \Rightarrow | \Rightarrow \Phi}\text{S5 rule}}{\Rightarrow \Box(\boxdot(\Box A \vee \Box B) \rightarrow A)| \Rightarrow \Phi}}{\Rightarrow \Box(\boxdot(\Box A \vee \Box B) \rightarrow A) \vee \Box(\boxdot(\Box A \vee \Box B) \rightarrow B)| \Rightarrow \Phi}}{\dfrac{\Rightarrow \Phi| \Rightarrow \Phi}{\Rightarrow \Phi}}$$

3.1.4 Deductive equivalence between Hilbert-style systems and hypersequent calculi

By using the following translation from hypersequents to formulas, we can prove that the hypersequents and Hilbert-style axiom systems are deductively equivalent.

To show this for monomodal cases, we use the following embedding from hypersequents to formulas.

$$\mathcal{I}(\Gamma_1 \Rightarrow \Delta_1| \dots |\Gamma_n \Rightarrow \Delta_n) = \Box(\bigwedge \Gamma_1 \rightarrow \bigvee \Delta_1) \vee \dots \vee \Box(\bigwedge \Gamma_n \rightarrow \bigvee \Delta_n)$$

In these combined systems, hypersequents are understood by an embedding $\mathcal{I}$, which is defined as follows.

$$\mathcal{I}(\Gamma_1 \Rightarrow \Delta_1| \dots |\Gamma_n \Rightarrow \Delta_n) = \boxdot(\bigwedge \Gamma_1 \rightarrow \bigvee \Delta_1) \vee \dots \vee \boxdot(\bigwedge \Gamma_n \rightarrow \bigvee \Delta_n).$$

Note that we do not use $\Box$ but we use $\boxdot$.[11]

---

[11] This is not an arbitrary choice. Apparently, there is no meaningful way of defining

In the following, L stands for one of the logics discussed above and HL stands for the hypersequent calculus for each of these.

Theorem 3.1.1 (Deductive Equivalence) *$\mathit{HL}\vdash \Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_n \Rightarrow \Delta_n$ if and only if $\mathit{L}\vdash \Box(\bigwedge\Gamma_1 \to \bigvee\Delta_1)\vee\cdots\vee\Box(\bigwedge\Gamma_n \to \bigvee\Delta_n)$.*

In the case of combined logics, we have $\mathit{HL}\vdash \Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_n \Rightarrow \Delta_n$ if and only if $L\vdash \boxdot(\bigwedge\Gamma_1 \to \bigvee\Delta_1)\vee\cdots\vee\boxdot(\bigwedge\Gamma_n \to \bigvee\Delta_n)$.

In the case of strict implication logics, we have the following.

For any $\mathit{L} \in \{$ $\mathit{SIS4}$, $\mathit{SIS4.3}$, $\mathit{SIS5}$ $\}$, $\mathit{HL}\vdash \Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_n \Rightarrow \Delta_n$ iff $\mathit{L}\vdash$ $(\bigwedge\Gamma_1 \longrightarrow \bigvee\Delta_1)\vee\cdots\vee(\bigwedge\Gamma_n \longrightarrow \bigvee\Delta_n)$, where "$\longrightarrow$" is a strict implication.

Proof Proof by induction on the length of the derivation for both directions.

$\Longleftarrow$) It suffices to prove that all the axioms of Hilbert-style systems are derivable and derivability is preserved under modus ponens. However, modus ponens, necessitation, a fortiori and adjunction can be taken to be a special case of cut, R$\Box$, R$\wedge$ and the combination of LW and R$\longrightarrow$. For instance, the last case is as follows.

$$\dfrac{\dfrac{\Rightarrow A}{B \Rightarrow A}}{\Rightarrow B \longrightarrow A}$$

Hence, we verify that all the axioms in Hilbert-style system are derivable. We focus on special modal axioms beyond S4, other cases being straightforward.

---

embedding into formulas by using $\Box$. This is a sharp contrast with the case of $\mathsf{S4LPN}$, where the embedding using $\Box$ for $\mathsf{S4}$ can be done smoothly. The problem consists in $R\boxdot$. (The problem is $\nvdash \Box A \to \Box \boxdot A$, which we need in order to prove soundness of the rule w.r.t. Hilbert-style system under the translation using $\Box$.)

1) .2 axiom:

$$\begin{array}{cl}
\varphi \Rightarrow \varphi & \\ \hline
\varphi, \neg\varphi \Rightarrow & \\ \hline\hline
\Box\varphi \Rightarrow, \Box\neg\varphi \Rightarrow & \\ \hline
\Box\varphi \Rightarrow |\Box\neg\varphi \Rightarrow & \text{restricted modal splitting} \\ \hline
\Box\varphi \Rightarrow | \Rightarrow \neg\Box\neg\varphi & \\ \hline
\Rightarrow \neg\Box\varphi | \Rightarrow \neg\Box\neg\varphi & \\ \hline
\Rightarrow \Box\neg\Box\varphi | \Rightarrow \Box\neg\Box\neg\varphi & \\ \hline
\neg\Box\neg\Box\varphi \Rightarrow | \Rightarrow \Box\neg\Box\neg\varphi & \\ \hline
\neg\Box\neg\Box\varphi \Rightarrow \Box\neg\Box\neg\varphi & \text{EC}
\end{array}$$

2) .3 Axiom

$$\begin{array}{cl}
\dfrac{\varphi \Rightarrow \varphi}{\Box\varphi \Rightarrow \varphi} \qquad \dfrac{\psi \Rightarrow \psi}{\Box\psi \Rightarrow \psi} & \\ \hline
\Box\varphi \Rightarrow \psi | \Box\psi \Rightarrow \varphi & \text{modal comm} \\ \hline
\Rightarrow \Box\varphi \to \psi | \Rightarrow \Box\psi \to \varphi & \\ \hline
\Rightarrow \Box(\Box\varphi \to \psi) | \Rightarrow \Box(\Box\psi \to \varphi) & \\ \hline
\Rightarrow \Box(\Box\varphi \to \psi) \vee \Box(\Box\psi \to \varphi) | \Rightarrow \Box(\Box\varphi \to \psi) \vee \Box(\Box\psi \to \varphi) & \\ \hline
\Rightarrow \Box(\Box\varphi \to \psi) \vee \Box(\Box\psi \to \varphi) & \text{EC}
\end{array}$$

3) Connecting axiom

$$\begin{array}{cl}
\varphi \Rightarrow \varphi & \\ \hline
\boxdot\varphi \Rightarrow \varphi & \\ \hline
\boxdot\varphi \Rightarrow \Box\varphi & \text{R}\Box \text{ in S4+L}
\end{array}$$

4) SIS4.3 (Notational convention : we write too long a line in a vertical manner. This applies to any case in the following in this chapter).

$$\begin{array}{cl}
\dfrac{A \Rightarrow A \qquad B \Rightarrow B}{A \longrightarrow B, A \Rightarrow B} \qquad \dfrac{C \Rightarrow C \qquad D \Rightarrow D}{C \longrightarrow D, C \Rightarrow D} & \\ \hline
C \longrightarrow D, A \Rightarrow B | A \longrightarrow B, C \Rightarrow D & \\ \hline
C \longrightarrow D \Rightarrow A \longrightarrow B | A \longrightarrow B \Rightarrow C \longrightarrow D & \text{si-comm} \\ \hline
\Rightarrow (C \longrightarrow D) \longrightarrow (A \longrightarrow B) | \Rightarrow (A \longrightarrow B) \longrightarrow (C \longrightarrow D) & \\ \hline
\Rightarrow ((C \longrightarrow D) \longrightarrow (A \longrightarrow B)) \vee ((A \longrightarrow B) \longrightarrow (C \longrightarrow A)) | & \\
\Rightarrow ((C \longrightarrow D) \longrightarrow (A \longrightarrow B)) \vee ((A \longrightarrow B) \longrightarrow (C \longrightarrow D)) & \\ \hline
\Rightarrow ((C \longrightarrow D) \longrightarrow (A \longrightarrow B)) \vee ((A \longrightarrow B) \longrightarrow (C \longrightarrow D)) & \text{EC}
\end{array}$$

5) SIS5

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A \Rightarrow A \qquad B \Rightarrow B}{A \longrightarrow B \Rightarrow A \longrightarrow B}\text{a few steps}}{A \longrightarrow B \Rightarrow \,|\, \Rightarrow A \longrightarrow B}\text{si-splitting}}{A \longrightarrow B \Rightarrow C| \Rightarrow A \longrightarrow B}}{\Rightarrow (A \longrightarrow B) \longrightarrow C| \Rightarrow A \longrightarrow B}}{\Rightarrow (A \longrightarrow B) \longrightarrow C, A \longrightarrow B| \Rightarrow (A \longrightarrow B) \longrightarrow C, A \longrightarrow B}}{\Rightarrow (A \longrightarrow B) \longrightarrow C, A \longrightarrow B}$$

$\Longleftarrow$) Conversely, we show that axioms in the hypersequent calculi and all the translated forms of rules in the hypersequent calculi are derivable. Since axioms and sequent rules are straightforward, we concentrate on external structural rules.

1) Modal splitting: By IH, $\vdash \bigvee G \vee \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Pi \to \bigvee \Theta)$. We derive $\vdash \bigvee G \vee \Box(\bigwedge \Box\Gamma \to \bot) \vee \Box(\bigwedge \Pi \to \bigvee \Theta)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Pi \to \bigvee \Theta) \to \Box(\Box \bigwedge \Gamma \wedge \bigwedge \Pi \to \bigvee \Theta)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Pi \to \bigvee \Theta) \to \Box(\Box \bigwedge \Gamma \to \bigwedge \Pi \to \bigvee \Theta)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Pi \to \bigvee \Theta) \to \Box\Box \bigwedge \Gamma \to \Box(\bigwedge \Pi \to \bigvee \Theta)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Pi \to \bigvee \Theta) \to \Box \bigwedge \Gamma \to \Box(\bigwedge \Pi \to \bigvee \Theta)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Pi \to \bigvee \Theta) \to \neg\Box \bigwedge \Gamma \vee \Box(\bigwedge \Pi \to \bigvee \Theta)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Pi \to \bigvee \Theta) \to \Box\neg\Box \bigwedge \Gamma \vee \Box(\bigwedge \Pi \to \bigvee \Theta)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Pi \to \bigvee \Theta) \to \Box(\Box \bigwedge \Gamma \to \bot) \vee \Box(\bigwedge \Pi \to \bigvee \Theta)$.

Hence, $\vdash \bigvee G \vee \Box(\Box \bigwedge \Gamma \to \bot) \vee \Box(\bigwedge \Pi \to \bigvee \Theta)$.

2) .2 rule: By IH, $\vdash \bigvee G \vee \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Box\Delta \to \bot)$.

On the other hand, by CPC, axioms and rules for $\mathsf{K4}$ ($\subseteq$ $\mathsf{S4.2}$) we can derive:

$\vdash (\bigwedge \Box\Gamma \wedge \bigwedge \Box\Delta \to \bot) \to (\Box \bigwedge \Gamma \wedge \Box \bigwedge \Delta \to \bot)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Box\Delta \to \bot) \to \Box(\Box \bigwedge \Delta \to \neg\Box \bigwedge \Gamma)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Box\Delta \to \bot) \to (\Box\Box \bigwedge \Delta \to \Box\neg\Box \bigwedge \Gamma)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Box\Delta \to \bot) \to (\Box \bigwedge \Delta \to \Box\neg\Box \bigwedge \Gamma)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Box\Delta \to \bot) \to (\neg\Box\neg\Box \bigwedge \Gamma \to \neg\Box \bigwedge \Delta)$.

$\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Box\Delta \to \bot) \to \Box\neg\Box\neg\Box \bigwedge \Gamma \to \Box\neg\Box \bigwedge \Delta$ $(*)$.

On the other hand, by S4.2 axiom and taking a substitution instance in which $\Box \bigwedge \Gamma$ is plugged into A, $\vdash \neg\Box\neg\Box\Box \bigwedge \Gamma \to \Box\neg\Box\neg\Box \bigwedge \Gamma$.

Also, $\vdash \Box \bigwedge \Gamma \to \Box\Box \bigwedge \Gamma$ yields $\vdash \neg\Box\neg\Box \bigwedge \Gamma \to \neg\Box\neg\Box\Box \bigwedge \Gamma$.

Hence, $\vdash \neg\Box\neg\Box \bigwedge \Gamma \to \Box\neg\Box\neg\Box \bigwedge \Gamma$.

Therefore, by $(*)$, $\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Box\Delta \to \bot) \to (\neg\Box\neg\Box \bigwedge \Gamma \to \Box\neg\Box \bigwedge \Delta)$.

This yields $\vdash \Box(\bigwedge \Box\Gamma \wedge \bigwedge \Box\Delta \to \bot) \to \Box(\bigwedge \Box\Gamma \to \bot) \vee \Box(\bigwedge \Box\Delta \to \bot)$.

Thus, $\vdash \bigvee G \vee \Box(\bigwedge \Box\Gamma \to \bot) \vee \Box(\bigwedge \Box\Delta \to \bot)$.

3) .3 Rule:

By IH, $\vdash \bigvee G \vee \Box(\bigwedge \Sigma \wedge \bigwedge \Box\Gamma \to \bigvee \Pi)$ and

$\vdash \bigvee G \vee \Box(\bigwedge \Theta \wedge \bigwedge \Box\Delta \to \bigvee \Lambda)$.

By normality, 4, and propositional logic,

$\vdash \Box(\bigwedge \Sigma \wedge \bigwedge \Box\Gamma \to \bigvee \Pi) \to (\Box \bigwedge \Delta \to (\Box \bigwedge \Gamma \to \Box(\bigwedge \Sigma \to \bigvee \Pi))$

$\vdash \Box(\bigwedge \Sigma \wedge \bigwedge \Box\Gamma \to \bigvee \Pi) \to ((\Box \bigwedge \Delta \to \Box \bigwedge \Gamma) \to (\Box \bigwedge \Delta \to \Box(\bigwedge \Sigma \to \bigvee \Pi)))$

$\vdash \Box(\bigwedge \Sigma \wedge \bigwedge \Box\Gamma \to \bigvee \Pi) \to ((\Box \bigwedge \Delta \to \Box \bigwedge \Gamma) \to (\Box \bigwedge \Delta \to (\bigwedge \Sigma \to \bigvee \Pi)))$

$\vdash \Box(\bigwedge \Sigma \wedge \bigwedge \Box\Gamma \to \bigvee \Pi) \to (\Box(\Box \bigwedge \Delta \to \Box \bigwedge \Gamma) \to (\Box(\bigwedge \Sigma \wedge \Box \bigwedge \Delta) \to \bigvee \Pi))$

Similarly, $\vdash \Box(\bigwedge \Theta \wedge \bigwedge \Box\Delta \to \bigvee \Lambda) \to (\Box(\Box \bigwedge \Gamma \to \Box \bigwedge \Delta) \to (\Box(\bigwedge \Theta \wedge \Box \bigwedge \Gamma) \to \bigvee \Lambda))$

Hence, $\vdash \bigvee G \vee (\Box(\Box \bigwedge \Delta \to \Box \bigwedge \Gamma) \to (\Box(\bigwedge \Sigma \wedge \Box \bigwedge \Delta) \to \bigvee \Pi))$ and $\vdash \bigvee G \vee (\Box(\Box \bigwedge \Gamma \to \Box \bigwedge \Delta) \to (\Box(\bigwedge \Theta \wedge \Box \bigwedge \Gamma) \to \bigvee \Lambda))$.

By .3, we can derive $\vdash \Box(\Box \bigwedge \Delta \to \Box \bigwedge \Gamma) \vee \Box(\Box \bigwedge \Gamma \to \Box \bigwedge \Delta)$. By propositional logic, $\bigvee G \vee \Box(\bigwedge \Sigma \wedge \Box \bigwedge \Delta) \to \bigvee \Pi) \vee \Box(\bigwedge \Theta \wedge \Box \bigwedge \Gamma) \to \bigvee \Lambda)$.

4) $R\Box$ in combined logics: By IH, S4 + L$\vdash \bigvee G \vee \boxdot(\bigwedge \Box\Gamma \wedge \bigwedge \boxdot\Delta \to \varphi)$

S4 + L$\vdash (\bigwedge \Box\Gamma \wedge \bigwedge \boxdot\Delta \to \varphi) \to (\Box \bigwedge \Gamma \wedge \boxdot \bigwedge \Delta \to \varphi)$.

S4 + L$\vdash \boxdot(\bigwedge \Box\Gamma \wedge \bigwedge \boxdot\Delta \to \varphi) \to \Box(\Box \bigwedge \Gamma \wedge \boxdot \bigwedge \Delta \to \varphi)$.

S4 + L$\vdash \boxdot(\bigwedge \Box\Gamma \wedge \bigwedge \boxdot\Delta \to \varphi) \to (\Box\Box \bigwedge \Gamma \wedge \Box \boxdot \bigwedge \Delta \to \Box\varphi)$.

S4 + L$\vdash \boxdot(\bigwedge \Box\Gamma \wedge \bigwedge \boxdot\Delta \to \varphi) \to (\Box \bigwedge \Gamma \wedge \boxdot \bigwedge \Delta \to \Box\varphi)$ $(S4 + L \vdash \boxdot A \to \Box \boxdot A)$.

S4 + L$\vdash \boxdot(\bigwedge \Box\Gamma \wedge \bigwedge \boxdot\Delta \to \varphi) \to \boxdot(\Box \bigwedge \Gamma \wedge \boxdot \bigwedge \Delta \to \Box\varphi)$

Hence, we have S4 + L$\vdash \bigvee G \vee \bigvee \boxdot(\bigwedge \Box\Gamma \wedge \bigwedge \boxdot\Delta \to \Box\varphi)$

SIS4.3: By IH,

$\vdash \bigvee G \vee (\bigwedge \Gamma \wedge \bigwedge \Theta_{\to} \longrightarrow \bigvee \Delta)$ and $\vdash \bigvee G \vee (\bigwedge \Sigma \wedge \bigwedge \Pi_{\to} \longrightarrow \bigvee \Lambda)$.

By SIS4 (the form $\bigwedge \Theta_{\to}$ is crucial here),

$\vdash ((\bigwedge \Theta_{\to} \wedge \bigwedge \Gamma) \longrightarrow \bigvee \Delta) \longrightarrow (\bigwedge \Theta_{\to} \longrightarrow (\bigwedge \Gamma \longrightarrow \bigvee \Delta))$.

$\vdash (\bigwedge \Theta_{\to} \longrightarrow (\bigwedge \Gamma \longrightarrow \bigvee \Delta)) \longrightarrow ((\bigwedge \Pi_{\to} \longrightarrow \bigwedge \Theta_{\to}) \longrightarrow (\bigwedge \Pi_{\to} \longrightarrow$

$(\bigwedge \Gamma \longrightarrow \bigvee \Delta))$.

Similarly, $\vdash (\bigwedge \Pi_{\to} \longrightarrow (\bigwedge \Sigma \longrightarrow \bigvee \Lambda)) \longrightarrow ((\bigwedge \Theta_{\to} \longrightarrow \bigwedge \Pi_{\to}) \longrightarrow (\bigwedge \Theta_{\to} \longrightarrow (\bigwedge \Sigma \longrightarrow \bigvee \Lambda))$.

$\vdash (\bigwedge \Gamma \wedge \bigwedge \Theta_{\to} \longrightarrow \bigvee \Delta) \longrightarrow ((\bigwedge \Pi_{\to} \longrightarrow \bigwedge \Theta_{\to}) \longrightarrow (\bigwedge \Gamma \wedge \bigwedge \Pi_{\to} \longrightarrow \bigvee \Delta))$ and

$\vdash (\bigwedge \Sigma \wedge \bigwedge \Pi_{\to} \longrightarrow \bigvee \Lambda) \longrightarrow ((\bigwedge \Theta_{\to} \longrightarrow \bigwedge \Pi_{\to}) \longrightarrow (\bigwedge \Sigma \wedge \bigwedge \Theta_{\to} \longrightarrow \bigvee \Lambda))$.

From an instance of the axiom of $\mathsf{SIS4.3}$, we can derive $\vdash (\bigwedge \Pi_{\to} \longrightarrow \bigwedge \Theta_{\to}) \vee (\bigwedge \Theta_{\to} \longrightarrow \bigwedge \Pi_{\to})$.[12] By $\mathsf{SIS4}$, $\vdash \bigvee G \vee (\bigwedge \Gamma \wedge \bigwedge \Pi_{\to} \longrightarrow \bigvee \Delta) \vee (\bigwedge \Sigma \wedge \bigwedge \Theta_{\to} \longrightarrow \bigvee \Lambda))$.

$\mathsf{SIS5}$: By IH, $\vdash \bigvee G \vee (\bigwedge \Gamma_{\to} \wedge \bigwedge \Pi \longrightarrow \bigvee \Delta)$.

From this, it is derived that $\vdash \bigvee G \vee (\bigwedge \Gamma_{\to} \longrightarrow (\bigwedge \Pi \longrightarrow \bigvee \Delta))$.

The following statement can be derived[13] in $\mathsf{SIS5} \vdash (\bigwedge \Gamma_{\to} \longrightarrow \bot) \vee \bigwedge \Gamma_{\to}$

We can prove $\bigvee G \vee (\bigwedge \Gamma_{\to} \longrightarrow \bot) \vee (\bigwedge \Pi \longrightarrow \bigvee \Delta)$ as follows.

$\vdash \bigvee G \vee ((\bigwedge \Gamma_{\to} \longrightarrow \bot) \vee \bigwedge \Gamma_{\to})$ and $\vdash \bigvee G \vee (\bigwedge \Gamma_{\to} \longrightarrow (\bigwedge \Pi \longrightarrow \bigvee \Delta))$

By $\mathsf{SIS4}$, $\vdash (\bigvee G \vee (\bigwedge \Gamma_{\to} \longrightarrow \bot)) \vee (\bigwedge \Gamma_{\to} \wedge (\bigwedge \Gamma_{\to} \longrightarrow (\bigwedge \Pi \longrightarrow \bigvee \Delta)))$.

Since $\mathsf{SIS4} \vdash \bigwedge \Gamma_{\to} \wedge (\bigwedge \Gamma_{\to} \longrightarrow (\bigwedge \Pi \longrightarrow \bigvee \Delta)) \longrightarrow (\bigwedge \Pi \longrightarrow \bigvee \Delta)$, we can derive $\vdash \bigvee G \vee (\bigwedge \Gamma_{\to} \longrightarrow \bot) \vee (\bigwedge \Pi \longrightarrow \bigvee \Delta)$.

---

[12] Unfortunately, purely syntactic verification of this theorem is too tedious. We use a semantic argument and completeness of the Hilbert-style system to justify this statement.

[13] Since this is tedious, we verified it semantically and appealed to completeness of the Hilbert-style system in this case, too.

3.2 Cut-elimination for hypersequent calculi for modal logics

Here we prove cut-elimination for modal logics (and strict implication logics) presented above. We confine ourselves to the ones that use hypersequents since sequent cases are not new. Here some logics are not new, but the uniformity in the proof of cut-elimination in hypersequent calculi for modal logics including S4 is part of our point here.

We start from the following general definitions. These definitions are given in [35]. These are understood as the results of applying cut multiple times.

A marked hypersequent is a hypersequent with exactly one occurrence of a formula A distinguished, written $(H|\Gamma, \underline{A} \Rightarrow \Delta)$ or $(H|\Gamma \Rightarrow \underline{A}, \Delta)$. A marked rule instance is a rule instance with the principal formula, if there is one, marked.

In [35], sequents are multisets of formulas. In our case here, sequents are sets. (Hypersequents are multisets of sequents.) Hence, all the superscripts are either 0 or 1 (hence, if $\lambda - 1 = 1$, then $\lambda = 1$, too). However, unfortunately, this does not really simplify the notation (we still need some symbol such as λ or μ although it is either 0 or 1). Therefore, we keep using the general notation.

Suppose that G is a (possibly marked) hypersequent and H a marked hypersequent of the forms:

$$G = (\Gamma_1, [A]^{\lambda_1} \Rightarrow \Delta_1 | \ldots | \Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n) \text{ and } H = (H' | \Pi \Rightarrow \underline{A}, \Sigma)$$

where A does not occur unmarked in $\biguplus_{i=1}^{n} \Gamma_i$. Then $CUT(G, H)$ is the set containing, for all $0 \leq \mu_i \leq \lambda_i$ for $i = 1 \ldots n$:

$$H' | \Gamma_1, \Pi^{\mu_1}, [A]^{\lambda_1 - \mu_1} \Rightarrow \Sigma^{\mu_1}, \Delta_1 | \ldots | \Gamma_n, \Pi^{\mu_n}, [A]^{\lambda_n - \mu_n} \Rightarrow \Sigma^{\mu_n}, \Delta_n.$$

Similarly, suppose that A does not occur unmarked in $\biguplus_{i=1}^{n} \Delta_i$ with:

$$G = (\Gamma_1 \Rightarrow [A]^{\lambda_1}, \Delta_1 | \ldots | \Gamma_n \Rightarrow [A]^{\lambda_n}, \Delta_n) \text{ and } H = (H' | \Pi, \underline{A} \Rightarrow \Sigma)$$

Then $CUT(G, H)$ contains, for all $0 \leq \mu_i \leq \lambda_i$ for $i = 1 \ldots n$:

$$H' | \Gamma_1, \Pi^{\mu_1} \Rightarrow [A]^{\lambda_1 - \mu_1}, \Sigma^{\mu_1}, \Delta_1 | \ldots | \Gamma_n, \Pi^{\mu_n} \Rightarrow [A]^{\lambda_n - \mu_n}, \Sigma^{\mu_n}, \Delta_n.$$

A rule (r) is *substitutive* if for any:

1. marked instance $\dfrac{G_1 \ldots G_n}{G}$ of (r);
2. marked hypersequent H appropriate for (r);
3. $G' \in CUT(G, H)$,

there exist $G'_i \in CUT(G_i, H)$ for $i = 1 \ldots n$ such that $\dfrac{G'_1 \ldots G'_n}{G'}$ of (r)

The point of introducing the notion of substitutivity is to use a way of eliminating cut without using the double induction that Gentzen used in his original proof of cut-elimination. If substitutivity holds, we can avoid using the measure of cut-height, since we can move up applications of cut themselves until we meet a line where these formulas are introduced by operational rules. (Let us quote from [35]. They say, "substitutivity ensures that cuts over formulas that are not principal in the rule can be shifted upwards over the premises.") Hence, we can concentrate on the complexity of a cut formula as the measure of induction. This is useful to handle external contraction, which raises a problem for Gentzen's original double induction. This method can be applied to a more extensive class of logics than the ones which simultaneous cut can cover.

Let us note one notational convention. In the following, when a binary rule becomes too wide, we write the rule in a vertical manner.

Lemma 3.2.1 *All non-modal rules in these logics (all the structural rules and operational rules for* $\wedge$*,* $\vee$*,* $\rightarrow$*) are substitutive.*

Proof The proof is just checking that each such rule satisfies the property, which is a (tedious) routine.

For modal cases, there is a problem that Mints meets in [106]. Namely, an arbitrary substitution on the antecedent disturbs $R\Box$ rule, since the antecedent has to be modalized for the rule to be applied. However, due to this very feature of the rule (i.e., the antecedent has to be modalized), it suffices to consider only modalized hypersequent of the form $H|\Box\Sigma \Rightarrow$ in order to permute applications of other rules and $R\Box$. In particular, if all the non-modal rules are substitutive and modal rules are substitutive for modalized cases, then it is possible to apply permutations of rules until we reach a rule that introduce the cut formula. This is the essence of the idea of the proof of cut-elimination here.[14]

Here we give a few more definitions. The *length* $|d|$ of a derivation d is (the maximal number of applications of inference rules)+1 occurring on any branch of d. The complexity of a formula A is the number of occurrences of its connectives. The *cut rank* $\rho(d)$ of d is (the maximal complexity of cut formulas in d)+1. Note that $\rho(d) = 0$ if d is cut-free.

Lemma 3.2.2 *Let* L *be such that* L$\in$ {*S4, S4.2, S4.3, S5* } *or* $L =$ *S4* + L *where* L = *S4.2, S4.3, S5. Let* d_l *and* d_r *be derivations in the hypersequent calculus* HL *for* L *such that:*

(1) d_l *is a derivation of* $(G|\Gamma_1, [A]^{\lambda_1} \Rightarrow \Delta_1|\dots|\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n)$ *(All occurrences of* A *are made explicit. There is no occurrence in* G*.);*

(2) d_r *is a derivation of* $(H|\Sigma \Rightarrow A, \Pi)$*;*

(3) $\rho(d_l) \leq |A|$ *and* $\rho(d_r) \leq |A|$*;*

---

[14] This method suggests a reply to Hacking's argument ([80]) against modal logic as logic because of its non-locality. Substitutivity of non-modal rules is due to the local nature (or context-independence) of the rules. However, for modal logics that proof-theoretically behave well, the non-locality of modal rules can be nicely handled by this notion of substitutivity w.r.t. modalized sequents. That is why the non-locality of modal rules does not spoil cut-elimination for modal logics that we are discussing. This may not apply to some other cases of modal logics such as K4.3.

(4) A is a compound formula and d_r ends with either a right logical rule or a modal rule introducing A.

Then a derivation d of $(G|H|\Gamma_1, \Sigma^{\lambda_1} \Rightarrow \Delta_1, \Pi^{\lambda_1}| \dots |\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n})$ *with* $\rho(d) \leq |A|$ *can be constructed in* HL.

Proof By induction on $|d_l|$. Base case is the one in which d_l is an axiom. In this case, there is nothing to do. The conclusion immediately holds.

Otherwise, we have different cases depending on the last rule applied to d_l.

Case 1. The last rule is applied on only side sequents G. We apply IH and we apply the rule on G. Then we get the desired hypersequent with $\rho(d) \leq |A|$. (Note all components in which A occurs are inactive in this step.)

Case 2. The last rule is any non-modal rule that does not have A as the principal formula. Then by the above lemma, the claim follows by IH and an application of the pertinent rule. For instance, $d_1 \vdash G|\Gamma_1, [A]^{\lambda_1}, B \Rightarrow \Delta_1| \dots |\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n$. The last inference in d_l is $L\wedge$ and it looks as follows.

$$\frac{G|\Gamma_1, [A]^{\lambda_1}, B \Rightarrow \Delta_1| \dots |\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}{G|\Gamma_1, [A]^{\lambda_1}, B \wedge C \Rightarrow \Delta_1| \dots |\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}$$

By IH, we get $G|H|\Gamma_1, \Sigma^{\lambda_1}, B \Rightarrow \Delta_1, \Pi^{\lambda_1}| \dots |\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n}$.

$$\frac{G|H|\Gamma_1, \Sigma^{\lambda_1}, B \Rightarrow \Delta_1, \Pi^{\lambda_1}| \dots |\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n}}{G|H|\Gamma_1, \Sigma^{\lambda_1}, B \wedge C \Rightarrow \Delta_1, \Pi^{\lambda_1}| \dots |\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n}}$$

The lower hypersequent in this inference is the desired one. (This is the kind of inference guaranteed by the lemma 1).

Case 3. The last inference is an application of left introduction rule whose principal formula is A, and $A = B * C$ where $* \in \{\wedge, \vee, \rightarrow, \neg\}$. Then the last inference of d_l looks as follows.

$$\frac{G|\Gamma_1, [B \wedge C]^{\lambda_1 - 1}, B \Rightarrow \Delta_1| \dots |\Gamma_n, [B \wedge C]^{\lambda_n} \Rightarrow \Delta_n}{G|\Gamma_1, [B \wedge C]^{\lambda_1} \Rightarrow \Delta_1| \dots |\Gamma_n, [B \wedge C]^{\lambda_n} \Rightarrow \Delta_n}$$

By IH, we obtain $G|H|\Gamma_1, \Sigma^{\lambda_1 - 1}, B \Rightarrow \Delta_1, \Pi^{\lambda_1}| \dots |\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n}$ with cut rank $\leq |A|$. Then, the desired hypersequent $G|H|\Gamma_1, \Sigma^{\lambda_1} \Rightarrow \Delta_1, \Pi^{\lambda_1}| \dots |\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n}$is derived by using cut with $(H|\Sigma \Rightarrow B, \Pi)$, which is one of the premise of the last inference rule of d_r (also with EW and EC). Note that this can be extracted because of the condition (4). Observe that the cut formula of the cut used in the derivation just mentioned is such that $|B| + 1 \leq |A|$. Hence, it does not increase the cut rank of the entire derivation, which is still $\leq |A|$.

Case 4. The rule of the last inference is $L\Box$ (We defer $R\Box$ to bimodal case that we discuss later since this is just a special case of the latter).

Subcase 4.1. A is the principal formula and $A = \Box B$. The last inference of d_l looks as follows.

$$\frac{G|\Gamma_1, B, [\Box B]^{\lambda_1 - 1} \Rightarrow \Delta_1| \dots |\Gamma_n, [\Box B]^{\lambda_n} \Rightarrow \Delta_n}{G|\Gamma_1, [\Box B]^{\lambda_1} \Rightarrow \Delta_1| \dots |\Gamma_n, [\Box B]^{\lambda_n} \Rightarrow \Delta_n}$$

Since d_r ends as the condition (4) states, the last inference of d_r is $\dfrac{H|\Box\Sigma' \Rightarrow B}{H|\Box\Sigma' \Rightarrow \Box B}$ (where $\Sigma = \Box\Sigma'$ and $\Pi = \emptyset$).

By IH, we can get $G|H|\Gamma_1, B, (\Box\Sigma')^{\lambda_1 - 1} \Rightarrow \Delta_1| \dots |\Gamma_n, (\Box\Sigma')^{\lambda_n} \Rightarrow \Delta_n$. The derivation of it has cut rank with $\leq |A|$.

By using $H|\Box\Sigma' \Rightarrow B$, cut, and EW, we can derive the desired $G|H|\Gamma_1, (\Box\Sigma')^{\lambda_1} \Rightarrow \Delta_1|\ldots|\Gamma_n, (\Box\Sigma')^{\lambda_n} \Rightarrow \Delta_n$. Note that the cut rank of the entire derivation is $\leq |A|$.

Subcase 4.2. The rule of the last inference of d_l is $L\Box$ and the principal formula is not A. Then the last inference of d_l looks as follows.

$$\frac{G|\Gamma_1, B, [A]^{\lambda_1} \Rightarrow \Delta_1|\ldots|\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}{G|\Gamma_1, \Box B, [A]^{\lambda_1} \Rightarrow \Delta_1|\ldots|\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}$$

By IH, $G|H|\Gamma_1, B, \Sigma^{\lambda_1} \Rightarrow \Delta_1, \Pi^{\lambda_1}|\ldots|\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n}$. The derivation of it has cut rank $\leq |A|$. Apply $L\Box$, we get $G|H|\Gamma_1, \Box B, \Sigma^{\lambda_1} \Rightarrow \Delta_1, \Pi^{\lambda_1}|\ldots|\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n}$. This is the desired hypersequent and its derivation has cut rank $\leq |A|$.

Case 5. The rule of the last inference in d_l is modal splitting.

Subcase 5.1. $A = \Box B$. Then the last inference of d_l looks as follows.

$$\frac{G|\Gamma_1', \Box\Theta, [\Box B]^{\lambda_1} \Rightarrow \Delta_1|\ldots|\Gamma_n, [\Box B]^{\lambda_n} \Rightarrow \Delta_n}{G|\Box\Theta, [\Box B]^{\lambda} \Rightarrow |\Gamma_1', [\Box B]^{\lambda_1-\lambda} \Rightarrow \Delta_1|\ldots|\Gamma_n, [\Box B]^{\lambda_n} \Rightarrow \Delta_n}$$

By IH, $G|H|\Gamma_1', \Box\Theta, (\Box\Sigma')^{\lambda_1} \Rightarrow \Delta_1|\ldots|\Gamma_n, (\Box\Sigma')^{\lambda_n} \Rightarrow \Delta_n$. The derivation of it has cut rank $\leq |A|$. Applying modal splitting, we get

$G|H|\Box\Theta, (\Box\Sigma')^{\lambda} \Rightarrow |\Gamma_1, (\Box\Sigma')^{\lambda_1-\lambda} \Rightarrow \Delta_1|\ldots|\Gamma_n, (\Box\Sigma')^{\lambda_n} \Rightarrow \Delta_n$.

This is the desired hypersequent and its derivation has cut rank $\leq |A|$.

Subcase 5.2. A is not of the form $\Box B$. The last inference of d_l looks as follows.

$$\frac{G|\Gamma_1', \Box\Theta, [A]^{\lambda_1} \Rightarrow \Delta_1|\ldots|\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}{G|\Box\Theta \Rightarrow |\Gamma_1', [A]^{\lambda_1} \Rightarrow \Delta_1|\ldots|\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}$$

Here $d_r \vdash H|\Sigma \Rightarrow A, \Pi$ (substitution is not restricted to modal substitution.)

By IH, $G|H|\Gamma_1, \Box\Theta, \Sigma^{\lambda_1} \Rightarrow \Pi^{\lambda_1}, \Delta_1|\ldots|\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Pi^{\lambda_n}, \Delta_n$. The derivation of it has cut rank $\leq |A|$. Applying modal splitting, we get

$G|H|\Box\Theta \Rightarrow |\Gamma_1, \Sigma^{\lambda_1} \Rightarrow \Pi^{\lambda_1}, \Delta_1|\ldots|\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Pi^{\lambda_n}, \Delta_n$.

This is the desired hypersequent and its derivation has cut rank $\leq |A|$.

Case 5′. The rule of the last inference of d_l is restricted modal splitting. It has to be the case that $A = \Box B$ due to the form of the premise of restricted modalized splitting. (A must be modalized if it occurs on the antecedent of the pertinent sequent $\Gamma \Rightarrow \Delta$.) Then the last inference of d_l looks as follows.

$$\frac{G|\Box\Xi, \Box\Theta, [\Box B]^{\lambda_1} \Rightarrow |\ldots|\Gamma_n, [\Box B]^{\lambda_n} \Rightarrow \Delta_n}{G|\Box\Xi, [\Box B]^{\lambda} \Rightarrow |\Box\Theta, [\Box B]^{\lambda_1-\lambda} \Rightarrow |\ldots|\Box\Gamma_n, [\Box B]^{\lambda_n} \Rightarrow \Delta_n}$$

By IH, $G|H|\Box\Xi, \Box\Theta, (\Box\Sigma')^{\lambda_1} \Rightarrow \Delta_1|\ldots|\Gamma_n, (\Box\Sigma')^{\lambda_n} \Rightarrow \Delta_n$. The derivation of it has cut rank $\leq |A|$. Applying restricted modal left splitting, we get

$G|H|\Box\Xi, (\Box\Sigma')^{\lambda} \Rightarrow |\Box\Theta, (\Box\Sigma')^{\lambda_1-\lambda} \Rightarrow \Delta_1|\ldots|\Gamma_n, (\Box\Sigma')^{\lambda_n} \Rightarrow \Delta_n$. This is the desired hypersequent and its derivation has cut rank $\leq |A|$.

Case 6. The rule of the last inference of d_l is modal comm.

Subcase 6.1. $A = \Box B$. The inference is as follows. Here $(\lambda_1, \lambda_1') = (1, 0)$, $(0, 1)$, $(1, 1)$ or $(0, 0)$.

$$\frac{\begin{array}{c}G|\Theta_1, \Box\Theta_2, [\Box B]^{\lambda_1} \Rightarrow \Delta|\ldots|\Gamma_n, [\Box B]^{\lambda_n} \Rightarrow \Delta_n \\ G|\Xi_1, \Box\Xi_2, [\Box B]^{\lambda_1'} \Rightarrow \Lambda|\ldots|\Gamma_n, [\Box B]^{\lambda_n} \Rightarrow \Delta_n\end{array}}{G|\Theta_1, \Box\Xi_2, [\Box B]^{\lambda_1'} \Rightarrow \Delta|\Xi_1, \Box\Theta_2, [\Box B]^{\lambda_1} \Rightarrow \Lambda|\ldots|\Box\Gamma_n, [\Box B]^{\lambda_n} \Rightarrow \Delta_n}$$

In this case $d_r \vdash H|\Box\Sigma' \Rightarrow \Box B$. The substitution is a modal substitution.

By IH, we can derive $G|H|\Theta_1, \Box\Theta_2, (\Box\Sigma')^{\lambda_1} \Rightarrow \Delta|\ldots|\Gamma_n, (\Box\Sigma')^{\lambda_n} \Rightarrow \Delta_n$ and $G|H|\Xi_1, \Box\Xi_2, (\Box\Sigma')^{{\lambda_1}'} \Rightarrow \Lambda|\ldots|\Gamma_n, (\Box\Sigma')^{\lambda_n} \Rightarrow \Delta_n$, and the derivation of it has cut rank $\leq |A|$. Applying modal comm, we get $G|H|\Theta_1, \Xi_1, \Box\Xi_2, (\Box\Sigma')^{{\lambda_1}'} \Rightarrow \Delta|\Theta_1, \Xi_1, \Box\Theta_2, (\Box\Sigma')^{\lambda_1} \Rightarrow \Lambda|\ldots|\Box\Gamma_n, (\Box\Sigma')^{\lambda_n} \Rightarrow \Delta_n$.

This is the desired hypersequent, and the derivation has the cut rank $\leq |A|$.

Subcase 6.2. A is not of the form $\Box B$. The last inference is as follows.

$$\frac{\begin{array}{c} G|\Theta_1, \Box\Theta_2, [A]^{\lambda_1} \Rightarrow \Delta|\ldots|\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n \\ G|\Xi_1, \Box\Xi_2, [A]^{{\lambda_1}'} \Rightarrow \Lambda|\ldots|\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n \end{array}}{G|\Theta_1, [A]^{\lambda_1}, \Box\Xi_2 \Rightarrow \Delta|\Xi_1, [A]^{{\lambda_1}'}, \Box\Theta_2 \Rightarrow \Lambda|\ldots|\Box\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}$$

Note that in this case d_r ends with the hypersequent $H|\Sigma \Rightarrow A, \Pi$.

By IH, we can derive $G|H|\Theta_1, \Box\Theta_2, \Sigma^{\lambda_1} \Rightarrow \Pi^{\lambda_1}, \Delta|\ldots|\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Pi^{\lambda_n}, \Delta_n$ and

$G|H|\Xi_1, \Box\Xi_2, \Sigma^{{\lambda_1}'} \Rightarrow \Pi^{{\lambda_1}'}, \Lambda|\ldots|\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Pi^{\lambda_n}, \Delta_n$, and the derivation of it has cut rank $\leq |A|$. Applying modal comm, we get $G|\Theta_1, \Sigma^{\lambda_1}, \Box\Xi_2 \Rightarrow \Pi^{\lambda_1}, \Delta|\Xi_1, \Sigma^{{\lambda_1}'}, \Box\Theta_2 \Rightarrow \Pi^{{\lambda_1}'}, \Lambda|\ldots|\Box\Gamma_n, [\Sigma]^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n}$.

This is the desired hypersequent, and the derivation has cut rank $\leq |A|$.

Case 7. d_l ends with $R\Box$.

Subcase 7.1 $A = \Box B$ The last inference of d_l is as follows.

$$\frac{G|\Box\Gamma_1', \boxdot\Theta, [\Box B]^{\lambda_1} \Rightarrow C|\ldots|\Gamma_n, [\Box B]^{\lambda_n} \Rightarrow \Delta_n}{G|\Box\Gamma_1', \boxdot\Theta, [\Box B]^{\lambda_1} \Rightarrow \Box C|\ldots|\Gamma_n, [\Box B]^{\lambda_n} \Rightarrow \Delta_n}$$

By the condition (4), the last inference of d_r is $\dfrac{H|\Box\Sigma_1, \boxdot\Sigma_2 \Rightarrow B}{H|\Box\Sigma_1, \boxdot\Sigma_2 \Rightarrow \Box B}$, and d_r ends with $d_l \vdash H|\Box\Sigma_1, \boxdot\Sigma_2 \Rightarrow \Box B$.

By IH, $G|H|\Box\Gamma'_1, \boxdot\Theta, [\Box\Sigma_1, \boxdot\Sigma_2]^{\lambda_1} \Rightarrow C|\ldots|\Gamma_n, [\Box\Sigma_1, \boxdot\Sigma_2]^{\lambda_n} \Rightarrow \Delta_n$. Applying $\Box R$ rule, we get
$G|H|\Box\Gamma'_1, \boxdot\Theta, [\Box\Sigma_1, \boxdot\Sigma_2]^{\lambda_1} \Rightarrow \Box C|\ldots|\Gamma_n, [\Box\Sigma_1, \boxdot\Sigma_2]^{\lambda_n} \Rightarrow \Delta_n$.

But this is the desired hypersequent.

Subcase 7.2 $A = \boxdot B$. The argument is similar to the above (by using $\boxdot\Sigma \Rightarrow$ for modal substitution).

Note: In the case of modal structural rules for combined logics, the same argument works as above (using $\boxdot$ instead of $\Box$), since $\Box$ is a S4 modality, which is inactive for modal structural rules.

Lemma 3.2.3 *Let L be such that $L \in \{$S4, S4.2, S4.3, S5 $\}$ or $L =$ S4 + L where L = S4.2, S4.3, S5. Let d_l and d_r be derivations in the hypersequent calculus HL for L such that:*

(1) d_l is a derivation of $(G|\Gamma, A \Rightarrow \Delta)$;

(2) d_r is a derivation of $(H|\Sigma_1 \Rightarrow [A]^{\lambda_1}, \Pi'_1|\ldots|\Sigma_n \Rightarrow [A]^{\lambda_n}, \Pi'_n)$;

(3) $\rho(d_l) \leq |A|$ and $\rho(d_r) \leq |A|$.

Then a derivation d of $(G|H|\Sigma_1, \Gamma^{\lambda_1} \Rightarrow \Pi'_1, \Delta^{\lambda_1}|\ldots|\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Pi'_n, \Delta^{\lambda_n})$ with $\rho(d) \leq |A|$ can be constructed in HL.

Proof By induction on $|d_r|$.

Base case: If the rule is the last inference is an axiom, then the statement obviously holds since there is no cut involved in d_r. (Hence the cut rank of $d =$ the cut rank of d_l. Thus, it is $\leq |A|$ by assumption.)

Inductive cases:

Case 1. the last inference is applied only to the side sequents H. Then we can apply IH, and since the pertinent component sequent is inactive, the statement immediately holds.

Case 2. The rule of the last inference is a non-modal rule where the marked occurrences of the cut formula A is not principal. (Here are subcases: L$\wedge$, L$\vee$, L$\rightarrow$, L$\neg$, and all non-modal structural rules.) These cases are uniformly done by IH and the above lemma stating that all non-modal rules are substitutive.

Case 3. The rule of the last inference a non-modal right introduction rule (an R rule) whose principal formula is A, and A is of the form $B * C$ where $* \in \{\wedge, \vee, \rightarrow\}$ (or $A = \neg B$). We take an subcase of $A = B \rightarrow C$ as a representative case for these and show it here. The last inference in d_r is as follows.

$$\frac{G|\Sigma_1, B \Rightarrow C, [B \rightarrow C]^{\lambda_1 - 1}, \Pi_1| \ldots |\Sigma_n \Rightarrow [B \rightarrow C]^{\lambda_n}, \Pi_n}{G|\Sigma_1 \Rightarrow [B \rightarrow C]^{\lambda_1}, \Pi_1| \ldots |\Sigma_n \Rightarrow [B \rightarrow C]^{\lambda_n}, \Pi_n}$$

By IH, we obtain $G|H|\Sigma_1, \Gamma^{\lambda_1 - 1}, B \Rightarrow C, \Delta^{\lambda_1 - 1}, \Pi_1| \ldots |\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n}$ with cut rank $\leq |A|$. Then, by applying $R \rightarrow$, we get $G|H|\Sigma_1, \Gamma^{\lambda_1 - 1} \Rightarrow B \rightarrow C, \Delta^{\lambda_1 - 1}, \Pi_1| \ldots |\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n}$ with cut rank $\leq |A|$. But then we can apply the previous lemma since we have only one occurrence of $A = B \rightarrow C$. Now the desired hypersequent $G|H|\Sigma_1, \Gamma^{\lambda_1} \Rightarrow \Delta^{\lambda_1}, \Pi_1'| \ldots |\Sigma_n, \Delta^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$ is derived. The cut rank of the entire derivation is $\leq |A|$.

Case 4. The rule of the last inference of d_r is L$\Box$.

In this case, A is not a principal formula. The last inference of d_l is as follows.

$$\frac{H|\Sigma_1, B \Rightarrow [A]^{\lambda_1}, \Pi_1| \ldots |\Sigma_n, \Rightarrow [A]^{\lambda_n}, \Pi_n}{H|\Sigma_1, \Box B, \Rightarrow [A]^{\lambda_1}, \Pi_1| \ldots |\Sigma_n, \Rightarrow [A]^{\lambda_n}, \Pi_n}$$

d_l ends with $G|\Gamma, A \Rightarrow \Delta$.

By IH, $G|H|\Sigma_1, B, \Gamma^{\lambda_1} \Rightarrow \Delta^{\lambda_1}, \Pi_1| \ldots |\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$ with the cut rank $\leq |A|$. Applying the L$\Box$ rule, we get $G|H|\Sigma_1, \Box B, \Gamma^{\lambda_1} \Rightarrow \Delta^{\lambda_1}, \Pi_1| \ldots$ $|\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$. This is the desired hypersequent. The cut rank of the derivation is $\leq |A|$.

Case 5. The rule of the last inference of d_r is $R\Box$. In this case, A has to be the principal formula of the form $\Box B$. The last inference of d_r is as follows.

$$\frac{H|\Box\Theta \Rightarrow B| \ldots |\Sigma_n \Rightarrow [\Box B]^{\lambda_n}, \Pi_n}{H|\Box\Theta \Rightarrow \Box B| \ldots |\Sigma_n \Rightarrow [\Box B]^{\lambda_n}, \Pi_n}$$

d_l ends with $G|\Gamma, \Box B \Rightarrow \Delta$.

By IH, $G|H|\Box\Theta \Rightarrow B| \ldots |\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$ with cut rank $\leq |A|$.

Applying $R\Box$, we get $H|\Box\Theta \Rightarrow \Box B| \ldots |\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$ with cut rank $\leq |A|$. By the previous lemma, the claim holds since we have only one occurrence of $\Box B$ in the entire hypersequent and satisfying the condition of application of the lemma.

Case 6. The rule of the last inference of d_r is modal splitting. Here we do not have to divide subcases, since A goes to only one place. Note that we have adopted a simpler form of modal splitting.

$$\frac{H|\Sigma_1', \Box\Theta \Rightarrow [A]^{\lambda_1}, \Pi_1'|\ldots|\Sigma_n \Rightarrow [A]^{\lambda_n}, \Pi_n}{H|\Box\Theta \Rightarrow |\Sigma_1' \Rightarrow [A]^{\lambda_1}, \Pi_1'|\ldots|\Sigma_n \Rightarrow [A]^{\lambda_n}, \Pi_n}$$

By IH, $G|H|\Sigma_1', \Box\Theta, \Gamma^{\lambda_1} \Rightarrow \Delta^{\lambda_1}, \Pi_1'|\ldots|\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$ is derivable with cut rank $\leq |A|$. Applying modal splitting, we can obtain the following hypersequent $G|H|\Box\Theta \Rightarrow |\Sigma_1', \Gamma^{\lambda_1} \Rightarrow \Delta^{\lambda_1}, \Pi_1'|\ldots|\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$. It is easy to check that this is the desired hypersequent, and the cut rank of the derivation is $\leq |A|$.

Case 6′. The rule of the last inference of d_r is restricted modal splitting. This case is similar to 6.

Case 7. The rule used in the last inference in d_r is modal comm. (Since Γ and Δ in $G|\Gamma, A \Rightarrow \Delta$ are arbitrary and the form of the premises does not enforce any particular formula for substitution, we do not have to divide subcases here.)

$$\frac{\begin{array}{c}H|\Theta_1, \Box\Theta_2 \Rightarrow [A]^{\lambda_1}, \Phi|\ldots|\Sigma_n \Rightarrow [A]^{\lambda_n} \Rightarrow \Pi_n \\ H|\Xi_1, \Box\Xi_2 \Rightarrow [A]^{\lambda_1'}, \Psi|\ldots|\Sigma_n \Rightarrow [A]^{\lambda_n}, \Rightarrow \Pi_n\end{array}}{H|\Theta_1, \Box\Xi_2 \Rightarrow [A]^{\lambda_1}, \Phi|\Xi_1, \Box\Theta_2 \Rightarrow [A]^{\lambda_1'}, \Psi|\ldots|\Box\Sigma_n \Rightarrow [A]^{\lambda_n}, \Pi_n}$$

By IH, we can derive $G|H|\Theta_1, \Box\Theta_2, \Gamma^{\lambda_1} \Rightarrow \Delta^{\lambda_1}, \Phi|\ldots|\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$ and

$G|H|\Xi_1, \Box\Xi_2, \Gamma^{\lambda_1'} \Rightarrow \Delta^{\lambda_1'}, \Psi|\ldots|\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$. Both derivations have cut rank $\leq |A|$. Applying modal comm, we get the following hypersequent.

$G|H|\Theta_1, \Box\Xi_2, \Gamma^{\lambda_1} \Rightarrow \Delta^{\lambda_1}, \Phi|\Xi_1, \Box\Theta_2, \Gamma^{\lambda_1'} \Rightarrow \Delta^{\lambda_1'}, \Psi|\ldots|\Box\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$. This is the desired hypersequent. The derivation of it has cut rank $\leq |A|$.

Case 8. d_r ends with $R\Box$. The only case is $A = \Box B$. $d_r \vdash \Gamma, \Box B \Rightarrow \Delta$.

$$\frac{H|\Box\Sigma_1, \boxdot\Theta \Rightarrow B|\ldots|\Sigma_n \Rightarrow [\Box B]^{\lambda_n}, \Pi_n}{H|\Box\Sigma_1, \boxdot\Theta \Rightarrow \Box B|\ldots|\Sigma_n \Rightarrow [\Box B]^{\lambda_n}, \Pi_n}$$

By IH, $G|H|\Box\Sigma_1, \boxdot\Theta \Rightarrow B|\ldots|\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$

Applying $R\Box$, we get $G|H|\Box\Sigma_1, \boxdot\Theta \Rightarrow \Box B|\ldots|\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Pi_n$.

But as in the case of $R\Box$ for monomodal logic, we have only one occurrence of $\Box B$ in the hypersequent. Hence, we can apply the above Lemma to show that we can further transform the derivation so that the cut rank of the derivation is $|A|$.

Theorem 3.2.4 (Cut-elimination) *For any modal logic* L *that in the set given above, cut-elimination holds for* HL.

Proof Let d be a derivation in a hypersequent calculus for a logic in L, with $\rho(d) > 0$. The proof proceeds by the double induction on $(\rho(d), n\rho(d))$, where $n\rho(d)$ is the number of application of cut in d with cut rank $\rho(d)$. Consider an uppermost application of cut in d with cut rank $\rho(d)$. By applying the last lemma to its premises. Either $\rho(d)$ or $n\rho(d)$ decreases. Then we can apply IH.

Cut-elimination for strict implication logics SIS4, SIS4.3, and SIS5 can be proven essentially in the same manner. Let us state the cut-elimination theorem for these logics and give an outline of a proof.

Theorem 3.2.5 *Let* L *be such that* L$\in$ {*SIS4, SIS4.3, SIS5*}.

For any such L, *cut-elimination holds for* HL.

We first prove two lemmas, but we give an outline since proofs are similar.

Let d_l and d_r be derivations in the hypersequent calculus HL for L such that:

Lemma 3.2.6 *Let L be such that L∈ {SIS4, SIS4.3, SIS5}.*

Let d_l and d_r be derivations in the hypersequent calculus HL for L such that:

(1) d_l is a derivation of $(G|\Gamma_1, [A]^{\lambda_1} \Rightarrow \Delta_1|\ldots|\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n)$ (All occurrences of A are made explicit. There is no occurrence in G.);

(2) d_r is a derivation of $(H|\Sigma \Rightarrow A, \Pi)$;

(3) $\rho(d_l) \leq |A|$ and $\rho(d_r) \leq |A|$;

(4) A is a compound formula and d_r ends with either a right operational rule introducing A.

Then a derivation d of $(G|H|\Gamma_1, \Sigma^{\lambda_1} \Rightarrow \Delta_1, \Pi^{\lambda_1}|\ldots|\Gamma_n, \Sigma^{\lambda_n} \Rightarrow \Delta_n, \Pi^{\lambda_n})$ with $\rho(d) \leq |A|$ can be constructed in HL.

Proof (In the cases of $\wedge$, $\vee$, the rules are substitutive. Hence, we do not need to discuss the details.)

Case 1. The last inference of d_l is $L \longrightarrow$. This is the same as intuitionistic case. The proofs given in the case of modal logics are, of course, based on classical logic, but as far as this case is concerned, there is no difference.

Now we deal with external structural rules. Since all the cases are similar to modal ones, we show only one representative case of SIS4.3 (si-comm).

Case 2. The last inference of d_l is the SIS4.3 rule (si-comm).

Subcase 2.1. A is of the form $B \longrightarrow C$. The last inference of d_l looks as follows. Here $(\lambda_1, \lambda_1') = (1,0)$, $(0,1)$, $(1,1)$ or $(0,0)$. Also, let us abbreviate $\Gamma_n, [B \longrightarrow C]^{\lambda_n} \Rightarrow \Delta_n$ as H_n

$$\frac{\begin{array}{c} G|\Phi, \Gamma_{\rightarrow}, [B \longrightarrow C]^{\lambda_1} \Rightarrow \Delta| \ldots |H_n \\ G|\Psi, \Theta_{\rightarrow}, [B \longrightarrow C]^{\lambda_1'} \Rightarrow \Lambda| \ldots |H_n \end{array}}{G|\Phi, \Theta_{\rightarrow}, [B \longrightarrow C]^{\lambda_1'} \Rightarrow \Delta|\Psi, \Gamma_{\rightarrow}, [B \longrightarrow C]^{\lambda_1} \Rightarrow \Lambda| \ldots |H_n}$$

By IH, we have $G|H|\Phi, \Gamma_{\rightarrow}, (\Sigma_{\rightarrow})^{\lambda_1} \Rightarrow \Delta| \ldots |\Gamma_n, (\Sigma_{\rightarrow})^{\lambda_n} \Rightarrow \Delta_n$ and $G|H|\Psi, \Theta_{\rightarrow}, (\Sigma_{\rightarrow})^{\lambda_1'} \Rightarrow \Lambda| \ldots |\Gamma_n, (\Sigma_{\rightarrow})^{\lambda_n} \Rightarrow \Delta_n$. Applying the rule, we can get $G|\Phi, \Theta_{\rightarrow}, (\Sigma_{\rightarrow})^{\lambda_1'} \Rightarrow \Delta|\Psi, \Gamma_{\rightarrow}, (\Sigma_{\rightarrow})^{\lambda_1} \Rightarrow \Lambda| \ldots |\Gamma_n, (\Sigma_{\rightarrow})^{\lambda_n} \Rightarrow \Delta_n$. But this is the desired hypersequent.

Subcase 2.2. A is not of the form $B \longrightarrow C$. The last inference of d_l looks as follows.

$$\frac{\begin{array}{c} G|\Phi, \Gamma_{\rightarrow}, [A]^{\lambda_1} \Rightarrow \Delta| \ldots |\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n \\ G|\Psi, \Theta_{\rightarrow}, [A]^{\lambda_1'} \Rightarrow \Lambda| \ldots |\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n \end{array}}{G|\Phi, \Theta_{\rightarrow}, [A]^{\lambda_1} \Rightarrow \Delta|\Psi, \Gamma_{\rightarrow}, [A]^{\lambda_1'} \Rightarrow \Lambda| \ldots |\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}$$

In this case, d_r ends with the hypersequent $H|\Sigma \Rightarrow A, \Pi$.

By IH, we get the hypersequents $G|H|\Phi, \Gamma_{\rightarrow}, (\Sigma)^{\lambda_1} \Rightarrow (\Pi)^{\lambda_1}, \Delta| \ldots |\Gamma_n, (\Sigma)^{\lambda_n} \Rightarrow \Delta_n$ and $G|H|\Psi, \Theta_{\rightarrow}, (\Sigma)^{\lambda_1'} \Rightarrow (\Pi)^{\lambda_1}, \Lambda| \ldots |\Gamma_n, (\Sigma)^{\lambda_n} \Rightarrow (\Pi)^{\lambda_n} \Rightarrow \Delta_n$.

Applying the rule, we can get the following hypersequent. $G|H|\Phi, \Theta_{\rightarrow}, (\Sigma)^{\lambda_1} \Rightarrow (\Pi)^{\lambda_1}, \Delta|\Psi, \Gamma_{\rightarrow}, (\Sigma)^{\lambda_1'} \Rightarrow (\Pi)^{\lambda_1}, \Lambda| \ldots |\Gamma_n, (\Sigma)^{\lambda_n} \Rightarrow \Delta_n$

But this is the desired hypersequent.

Lemma 3.2.7 *Let L be such that L∈ {SIS4, SIS4.3, SIS5}.*

Let d_l and d_r be derivations in the hypersequent calculus HL for Lsuch that:

(1) d_l *is a derivation of* $(G|\Gamma, A \Rightarrow \Delta)$*;*

(2) d_r *is a derivation of* $(H|\Sigma_1 \Rightarrow [A]^{\lambda_1}, \Pi'_1|\ldots|\Sigma_n \Rightarrow [A]^{\lambda_n}, \Pi'_n)$*;*

(3) $\rho(d_l) \leq |A|$ *and* $\rho(d_r) \leq |A|$*.*

Then a derivation d of $(G|H|\Sigma_1, \Gamma^{\lambda_1} \Rightarrow \Pi'_1, \Delta^{\lambda_1}|\ldots|\Sigma_n, \Gamma^{\lambda_n} \Rightarrow \Pi'_n, \Delta^{\lambda_n})$ *with* $\rho(d) \leq |A|$ *can be constructed in* HL*.*

Proof Case 1. The last inference of d_r is $L \longrightarrow$. Similar to classical case. There is no special condition.

Case 2. The last inference of d_r is $R \longrightarrow$. The last inference of d_r looks as follows.

$$\frac{H|\Sigma_{\rightarrow}, B \Rightarrow C|\ldots|\Sigma_n \Rightarrow [B \longrightarrow C]^{\lambda_n}, \Pi}{H|\Sigma_{\rightarrow} \Rightarrow B \longrightarrow C|\ldots|\Sigma_n \Rightarrow [B \longrightarrow C]^{\lambda_n}, \Pi}$$

d_l ends with a hypersequent of the general form $G|\Gamma \Rightarrow \Delta$.

By IH, we can derive the hypersequent $G|H|\Sigma_{\rightarrow}, B \Rightarrow C|\ldots|\Sigma_n, (\Gamma)^{\lambda_n} \Rightarrow (\Delta)^{\lambda_n}, \Pi$. Applying $R \longrightarrow$, we get $G|H|\Sigma_{\rightarrow} \Rightarrow B \rightarrow C|\ldots|\Sigma_n, (\Gamma)^{\lambda_n} \Rightarrow (\Delta)^{\lambda_n}, \Pi$. Since the pertinent $B \longrightarrow C$ is the only occurrence of $B \longrightarrow C$, we can apply the previous lemma. Then, the statement holds.

We have the following further cases.

Case 3. The last inference in d_r is SIS4.3 rule (si-com).

Case 4. The last inference in d_r is SIS5 rule (si-splitting).

However, in these external structural rules, all the special parts that involve strict implications and need special care are on the antecedent. We do not have any condition to be put on the relevant formula A. The situation is the same as these in the case of modal logics, and proofs are straightforward.

Proof (of the cut-elimination theorem)

The argument is essentially the same as in the case of modal logics. ☒

Notes on intuitionistic variants of modal logics extending S4:

Although we do not intend to give a systematic treatment of intuitionistic analogues of modal logics, for some modal logics among these, we can readily find out their intuitionistic analogues in the literature. Here are such cases.

classical	intuitionistic
$\mathsf{S4}$	$\mathsf{IS4}_\Box = \mathrm{IPC} + \mathsf{KT4} + \text{Necessitation}$
$\mathsf{S4.3}$	$\mathsf{IS4.3}_\Box = \mathsf{IS4}_\Box + \Box(\Box A \to B) \vee \Box(\Box B \to A)$
S5	$\mathsf{GIp} = \mathsf{IS4}_\Box + (\Box A \to \Box B) \to \Box(\Box A \to \Box B) + \Box A \vee \neg\Box A$

Let us note a few points.

1. Traditionally intuitionistic modal logics are formulated by using both $\Box$ and $\Diamond$, but the systems we gave are only $\Box$ fragment. $\mathsf{IS4}_\Box$ is the basic system here.

2. There are some subtleties in formulating intuitionistic modal logics. For instance, in classical setting, adding $(\Box A \to \Box B) \to \Box(\Box A \to \Box B)$ is sufficient to formulate a version of intuitionistic $\mathsf{S5}$. However, in intuitionistic setting, there are infinitely many different formulations in intuitionistic $\mathsf{S5}_\Box$.

The one given above is a very strong formulation given for certain specific purpose in global intuitionistic set theory ([33]). IS4.3 is given in [177] as one of the logics ($\Box$-fragments) for which the finite model property holds.

3. S4.2 does not seem to have an existing simple intuitionistic analogue in the literature.

4. Hilbert-style systems for these logics can be formulated by modifying the underlying logic of the classical case to intuitionistic logic.

5. Similarly, classical versions of hypersequent calculi for these logics can be turned into the intuitionistic version of the corresponding logics by restricting the number of formulas occurring in each sequent of a hypersequent.

It is easy to observe that all the steps in the lemmas for the cut-elimination theorem can be applied to these intuitionistic versions, since in the crucial modal cases rules are applied only to the antecedent. Hence, cut-elimination for the hypersequent calculi for these intuitionistic variants are corollaries of those for classical case.

3.3 Gödel embedding for superintuitionistic logics into modal logics

Before moving on to the issue of embedding, we add a brief note on the relationship among strict implication logics, superintuitionistic logics, and modal logics.

1. The relationship among these logics can be described as follows.

1.1. By adding $A \to (B \to A)$ in the Hilbert-style systems of the strict implication logics, the resulting logics become the corresponding superintuitionistic logics.

1.2. In hypersequent calculi, the difference of the two different kinds of logics can be encoded only in the feature of $R \to$. In strict implication logics, we restrict the antecedent of the pertinent sequent to $\Gamma_{\to}$. If we remove this restriction and allow arbitrary formulas in Γ, then we get the corresponding superintuitionistic logic.

1.3. Semantically, these differences rest on truth persistence. (Frame conditions are also slightly different.)

2. All the modal operators present in such modal logics as $\mathsf{S4}$, $\mathsf{S4.3}$, $\mathsf{S5}$ can be definable via strict implications by using a formula $\top \to A = \Box A$. (We conjecture that the case of $\mathsf{S4.2}$ holds for $\mathsf{S4.2}$.) In all systems, material implication is added. Cut-elimination for the entire system with $\Box$ and material implication guarantee that this addition is a conservative extension.[15]

3. In contrast to definability results mentioned in two, we have Gödel embedding of superintuitionistic logics into their modal counterparts. One application of cut-elimination of the modal logics discussed above ($\mathsf{S4}$, $\mathsf{S4.2}$, $\mathsf{S4.3}$, $\mathsf{S5}$) is a syntactic proof of soundness and faithfulness of Gödel embed-

---

[15] We do not verify this for strict implication logics directly here, except the $\mathsf{SIS4}$ case (see the appendix).

ding from (super)intuitionistic logics into their modal counterpart.

Here we discuss Gödel embedding of superintuitionistic logics (we call them the set of logics L = { intuitionistic logic, logic of weak excluded middle, Gödel-Dummett logic, classical logic } into modal logics discussed above (we call them ML (= { S4, S4.2, S4.3, S5 }) from a proof theoretic point of view. We prove soundness and faithfulness of the embedding.

There are several variants of so-called Gödel embedding commonly cited in the literature. We use Maehara's version ([100]) since this version is helpful for our purpose.

Definition 3.3.1 *The embedding is defined by induction as follows.*

1. $p^{\Box} := \Box p$; *2.* $\bot^{\Box} := \bot$
2. $(A \wedge B)^{\Box} := \Box(A^{\Box} \wedge B^{\Box})$
3. $(A \vee B)^{\Box} := \Box(A^{\Box} \vee B^{\Box})$
4. $(A \rightarrow B)^{\Box} := \Box(A^{\Box} \rightarrow B^{\Box})$

Note that all the formulas except $\bot$ that are in the set of images of the mapping are all modalized. We first state a proposition.

To make our argument simpler, we use both single-conclusion hypersequent calculi (we call them sHL) and multiple-conclusion hypersequent calculi (we call them mHL) for logics in L. The latter are analogous to Maehara's multiple-conclusion sequent calculus for intuitionistic logic. Such multiple-conclusion hypersequent calculi are obtained from single-conclusion versions as follows: 1. keep intuitionistic R→; 2. all other rules are formulated as multiple-conclusion versions including external structural rules characterizing logics.[16] The rules are presented in 2.5.1. The equivalence of the two calculi can be proven in a straightforward manner. (We use cut-free single-conclusion hypersequent calculi, but the multiple-conclusion ones do not have to be cut-free. In the following, HML^- and sHL^- mean cut-free HML and sHL, respectively.)

We use the notation for sets of formulas $\Gamma^\Box := B_1^\Box, \ldots, B_n^\Box$ and multisets of sequents $G^\Box := \Gamma_1^\Box \Rightarrow \Delta_1^\Box | \ldots | \Gamma_n^\Box \Rightarrow \Delta_n^\Box$.

Since $\bot^\Box = \bot$, without loss of generality, we assume that $\bot^\Box \notin \Gamma_i^\Box$ for any antecedent Γ_i in hypersequents occurring in a cut-free proof. $\bot^\Box \in \Gamma_i^\Box$ for some antecedent in a given hypersequent or $\bot^\Box \notin \Gamma_i^\Box$ for any antecedent in the hypersequent. Suppose $\bot^\Box = \bot \in \Gamma_1^\Box$ for $\Box\Gamma_1$ in a given hypersequent s.t. $\mathrm{sHL}^-\vdash G|\bot, \Gamma_1 \Rightarrow \varphi_1 | \ldots | \Gamma_n \Rightarrow \varphi_n$. In this case, $\mathsf{HML}\vdash G^\Box | \bot^\Box, \Gamma_1^\Box \Rightarrow \varphi_1^\Box | \ldots | \Gamma_n^\Box \Rightarrow \varphi_n^\Box$ trivially follows. Hence, it suffices to show the second case in which $\bot^\Box (= \bot) \notin \Gamma_i^\Box$.

Proposition 3.3.2 (Soundness) *If sHL*$^-\vdash G|\Gamma_1 \Rightarrow \varphi_1 | \ldots | \Gamma_n \Rightarrow \varphi_n$*, then* *HML*$\vdash G^\Box | \Gamma_1^\Box \Rightarrow \varphi_1^\Box | \ldots | \Gamma_n^\Box \Rightarrow \varphi_n^\Box$.

Proof Proof is by induction on the length of cut-free proofs of sHL. (Proofs

---

[16] One exception is the rule for logic of weak excluded middle, since this rule originally has the empty succedent.

vary depending on which characteristic external modal structural rule a system of ML has. We give a proof for each case of these rules.)

Case 1. Axioms

In the case of axiom of the form sHL$\vdash$ $p \Rightarrow p$, the claim holds as follows.

By the mapping, we get $p^{\Box} \Rightarrow p^{\Box}$, i.e. $\Box p \Rightarrow \Box p$. This is certainly provable. For $p \Rightarrow p$ implies $\Box p \Rightarrow p$, and then $\Box p \Rightarrow \Box p$ follows.

The case of the other form of the axiom sHL$\vdash$ $\bot \Rightarrow$ has $\bot$ on the antecedent. Hence, it can be handled as we discussed above. (Indeed, $\bot^{\Box} = \bot$. Hence, $\bot^{\Box} \Rightarrow$ immediately follows.)

Case 2. Operational rules. Most cases are straightforward. As representative cases, we show the cases of L$\rightarrow$ and $R \rightarrow$.

Subcase 2.1. L$\rightarrow$ $\dfrac{G|\Gamma \Rightarrow A \qquad G|B, \Pi \Rightarrow C}{G|\Gamma, \Pi, A \rightarrow B \Rightarrow C}$

By IH, HML$\vdash$ $G^{\Box}|\Gamma^{\Box} \Rightarrow A^{\Box}$ and HML$\vdash$ $G^{\Box}|B^{\Box}, \Pi^{\Box} \Rightarrow C^{\Box}$.

Then the image of the mapping $()^{\Box}$ is derived as follows.

$$\dfrac{\dfrac{\dfrac{G^{\Box}|\Gamma^{\Box} \Rightarrow A^{\Box} \qquad G^{\Box}|B^{\Box}, \Pi^{\Box} \Rightarrow C^{\Box}}{G^{\Box}|\Gamma^{\Box}, \Pi^{\Box}, A^{\Box} \rightarrow B^{\Box} \Rightarrow C^{\Box}}}{G^{\Box}|\Gamma^{\Box}, \Pi^{\Box}, \Box(A^{\Box} \rightarrow B^{\Box}) \Rightarrow C^{\Box}}\,\text{L}\Box}{G^{\Box}|\Gamma^{\Box}, \Pi^{\Box}, (A \rightarrow B)^{\Box} \Rightarrow C^{\Box}}$$

Subcase 2.2. $R \rightarrow$ $\dfrac{G|\Gamma, A \Rightarrow B}{G|\Gamma \Rightarrow A \rightarrow B}$ (It is the same as the intuitionistic $R \rightarrow$ rule.)

By IH, the translation of the premise of the rule $R\rightarrow$ in sHL, i.e. HML$\vdash$ $G^{\Box}|\Gamma^{\Box}, A^{\Box} \Rightarrow B^{\Box}$. Then we can have the following derivation. Note that since $\Gamma^{\Box}$ are already modalized, there is no need of applying cuts.

$$\cfrac{\cfrac{\cfrac{G^{\Box}|\Gamma^{\Box}, A^{\Box} \Rightarrow B^{\Box}}{G^{\Box}|\Gamma^{\Box} \Rightarrow A^{\Box} \rightarrow B^{\Box}}}{G^{\Box}|\Gamma^{\Box} \Rightarrow \Box(A^{\Box} \rightarrow B^{\Box})}\,\mathrm{R}\Box}{G^{\Box}|\Gamma^{\Box} \Rightarrow (A \rightarrow B)^{\Box}}$$

The lowest hypersequent is the desired hypersequent since this is the image of the mapping $()^{\Box}$ of the lower hypersequent of the inference in L.

Case 3. Structural rules

The case of the standard structural rules for the hypersequent calculus for intuitionistic propositional logic are all straightforward. Hence, we check only external structural rules characterizing logics properly extending intuitionistic logic.

Subcase 3.1. Splitting (L = classical logic)

$$\frac{G|\Gamma_1, \Gamma_2 \Rightarrow \varphi_1| \dots |\Gamma_n \Rightarrow \varphi_n}{G|\Gamma_1 \Rightarrow \varphi_1|\Gamma_2 \Rightarrow |\Gamma_n \Rightarrow \varphi_n}$$

By IH, HS5$\vdash$ $G^{\Box}|\Gamma_1^{\Box}, \Gamma_2^{\Box} \Rightarrow \Delta_1^{\Box}| \dots |\Gamma_n^{\Box} \Rightarrow \Delta_n^{\Box}$.

Since $\Gamma_2^{\Box}$ are modalized formulas, we do not have to do anything to make them modalized.

By applying modal splitting and cuts, we get the following inference.

$$\frac{G^{\Box}|\Gamma_1^{\Box}, \Gamma_2^{\Box} \Rightarrow \varphi_1^{\Box}| \dots |\Gamma_n^{\Box} \Rightarrow \varphi_n^{\Box}}{G^{\Box}|\Gamma_1^{\Box} \Rightarrow \varphi_1^{\Box}|\Gamma_2^{\Box} \Rightarrow |\Gamma_n^{\Box} \Rightarrow \varphi_n^{\Box}}$$

But the lower hypersequent is the image of the mapping $()^{\Box}$ applied to the lower hypesequent of the inference of classical splitting. Hence, the inference is preserved under this mapping.

Subcase 3.2. Restricted splitting (L is LWEM).

$$\frac{G|\Gamma_1, \Gamma_2 \Rightarrow |\ldots|\Gamma_n \Rightarrow \varphi_n}{G|\Gamma_1 \Rightarrow |\Gamma_2 \Rightarrow |\Gamma_n \Rightarrow \varphi_n}$$

This is simply a special case of the case of splitting where we have the empty set on the succedent. Hence the argument is essentially the same as above.

Subcase 3.3. Communication (L is Gödel-Dummett logic)

$$\frac{G|\Gamma_1, \Pi \Rightarrow \varphi_1|\ldots|\Gamma_n \Rightarrow \varphi_n \qquad G|\Gamma_2, \Sigma \Rightarrow \varphi_2|\ldots|\Gamma_n \Rightarrow \varphi_n}{G|\Gamma_1, \Sigma \Rightarrow \varphi_1|\Gamma_2, \Pi \Rightarrow \varphi_2|\ldots|\Gamma_n \Rightarrow \varphi_n}$$

By IH, we have as follows $\mathsf{HS4.3}\vdash G^{\Box}|\Gamma_1^{\Box}, \Pi^{\Box} \Rightarrow \varphi_1^{\Box}|\ldots|\Gamma_n^{\Box} \Rightarrow \varphi_n^{\Box}$ and $\mathsf{HS4.3}\vdash G^{\Box}|\Gamma_2^{\Box}, \Sigma^{\Box} \Rightarrow \varphi_2^{\Box}|\ldots|\Gamma_n^{\Box} \Rightarrow \varphi_n^{\Box}$. Note again that $\Pi^{\Box}$ and $\Sigma^{\Box}$ are modalized.

$$\frac{G^{\Box}|\Gamma_1^{\Box}, \Pi^{\Box} \Rightarrow \varphi_1^{\Box}|\ldots|\Gamma_n^{\Box} \Rightarrow \varphi_n^{\Box} \qquad G^{\Box}|\Gamma_2^{\Box}, \Sigma^{\Box} \Rightarrow \varphi_2^{\Box}|\ldots|\Gamma_n^{\Box} \Rightarrow \varphi_n^{\Box}}{G^{\Box}|\Gamma_1^{\Box}, \Sigma^{\Box} \Rightarrow \varphi_1^{\Box}|\Gamma_2^{\Box}, \Pi^{\Box} \Rightarrow \varphi_2^{\Box}|\ldots|\Gamma_n^{\Box} \Rightarrow \varphi_n^{\Box}}$$

It is easy to check that the lower hypersequent is the image of the mapping applied to the lowersequent of the original case of communication. Hence, the derivability is preserved under this mapping. This completes the proof of the proposition. ⊠

Theorem 3.3.3 (Faithfulness) *If HML⁻⊢* $G^{\Box}|\Gamma_1^{\Box} \Rightarrow \Delta_1^{\Box}|\ldots\ |\Gamma_n^{\Box} \Rightarrow \Delta_n^{\Box}$, *then mHL⊢* $G|\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_n \Rightarrow \Delta_n$.

Proof Faithfulness of the embedding $()^{\Box}$ is proven as follows.

The basic idea is simply to strip off all the modal operators from a given cut-free proof of HML in the assumption.

Case 1. HML = $\mathsf{HS5}$ and HL = classical logic.

The underlying logic of HS5 is classical logic, and there is no way that using a multiple-conclusion hypersequent and having the external structural rule of unconditional splitting for arbitrary formulas make the underlying logic stronger than classical logic. Therefore, stripping off all modal operators from the given proof in HS5 simply gives a proof in a multiple-conclusion hypersequent calculus with intuitionistic R→ and with additional rule splitting, which makes the system classical. This establishes that the variant of the embedding is faithful.

Case 2. HML = $\mathsf{HS4.2}$ or $\mathsf{HS4.3}$ and HL = LWEM or Gödel-Dummett logic (respectively). These cases require some additional care to prove the faithfulness of the embedding. Inferences that matter are external modal structural rules or R→. (Other cases are straightforward.)

Case 2.1. External modal structural rules.

We need to guarantee that after stripping off the modal operators, applying external structural rules applied in the hypersequent calculi for the logics does give valid inference in the desired logics. However, these are easy observations.

2.1.1. Suppose we have restricted modal splitting.

$$\frac{G^{\Box}|\Gamma_1^{\Box},\Gamma_2^{\Box}\Rightarrow|\dots|\Gamma_n^{\Box}\Rightarrow\Delta_n^{\Box}}{G^{\Box}|\Gamma_1^{\Box}\Rightarrow|\Gamma_2^{\Box}\Rightarrow|\Gamma_n^{\Box}\Rightarrow\Delta_n^{\Box}}$$

If we strip off all the modal operators, then we have the following inference.

$$\frac{G|\Gamma_1,\Gamma_2\Rightarrow|\dots|\Gamma_n\Rightarrow\Delta_n}{G|\Gamma_1\Rightarrow|\Gamma_2\Rightarrow|\Gamma_n\Rightarrow\Delta_n}$$

This is exactly the same inference as the restrict weakening valid in LWEM. Hence, this case satisfies the required property.

2.1.2. Suppose we have modal communication.

$$\frac{G^{\Box}|\Gamma_1^{\Box},\Pi^{\Box}\Rightarrow\Delta_1^{\Box}|\dots|\Gamma_n^{\Box}\Rightarrow\Delta_n^{\Box}\qquad G^{\Box}|\Gamma_2^{\Box},\Sigma^{\Box}\Rightarrow\Delta_2^{\Box}|\dots|\Gamma_n^{\Box}\Rightarrow\Delta_n^{\Box}}{G^{\Box}|\Gamma_1^{\Box},\Sigma^{\Box}\Rightarrow\Delta_1^{\Box}|\Gamma_2^{\Box},\Pi^{\Box}\Rightarrow\Delta_2^{\Box}|\dots|\Gamma_n^{\Box}\Rightarrow\Delta_n^{\Box}}$$

If we strip off all the modal operators, then we have the following inference.

$$\frac{G|\Gamma_1,\Pi\Rightarrow\Delta_1|\dots|\Gamma_n\Rightarrow\Delta_n\qquad G|\Gamma_2,\Sigma\Rightarrow\Delta_2|\dots|\Gamma_n\Rightarrow\Delta_n}{G|\Gamma_1,\Sigma\Rightarrow\Delta_1|\Gamma_2,\Pi\Rightarrow\Delta_2|\dots|\Gamma_n\Rightarrow\Delta_n}$$

Case 2.2. R$\to$. We show that all the applications R$\to$ can be restricted to a single-conclusion version. Then, after all the $\Box$'s are stripped off, we have intuitionistically valid $R\to$.

First observe that all the occurrences of $A\to B$ statements are in the scope of $\Box$ in the range of the mapping $()^{\Box}$ in $G^{\Box}|\Gamma^{\Box}\Rightarrow\Delta^{\Box}$.

Also, all the positive occurrences of $\Box$-ed formulas in a cut-free proof of $G^{\Box}|\Gamma^{\Box}\Rightarrow\Delta^{\Box}$ can be (in general) introduced in one of the following three ways.

1. Axiom; 2. Weakening; 3. R$\Box$

However, in our case, 1 is out, since we confine ourselves only to proof systems in which only atomic formulas are used in axioms. In case 2, it is clear that if we strip off modal operators from everywhere in a proof figure, the inference of the case 2 is still (intuitionistically) valid in the multiple-conclusion hypersequent calculus.

The only remaining case is that $\Box(A \to B)$ introduced via R$\Box$ from $A \to B$. Now positive occurrences of $A \to B$ can be introduced via $R \to$ or weakening (IW or EW). But the only case that matters is again the case where this is introduced by $R \to$. Then, to show that all the applications of R$\to$ in HML^- can be restricted to a single-conclusion case, it suffices to show that all the cases of positive occurrences of a formula of the form $\Box(A \to B)$ are introduced in such a way that the introduction of $A \to B$ in HML^- can be restricted to the single-conclusion version of R$\to$ and then R$\Box$ is applied (since all the cases that matter here are the ones in the range of the mapping $()^\Box$.) Then, it has to be introduced as follows.

$$\cfrac{\cfrac{G^\Box | \Box\Gamma, A^\Box \Rightarrow B^\Box}{G^\Box | \Box\Gamma' \Rightarrow A^\Box \to B^\Box}\text{some intermediate steps}}{G^\Box | \Box\Gamma'' \Rightarrow \Box(A^\Box \to B^\Box)}\text{some intermediate steps}$$

In modal logic, in general, there may be some applications of R$\to$ in which there are more than one formula on the succedent. (Such cases are not initially excluded.)

However, (1) there is no way of applying R$\Box$ to a sequent in which we have more than one formula on the succedent; (2) we want to derive $\Box(A^{\Box} \to B^{\Box})$, which is in the range of the mapping $()^{\Box}$, and applying R$\Box$ to the very formula $A^{\Box} \to B^{\Box}$ is the only way of doing that.

For these two reasons, we must have a sequent in which $A^{\Box} \to B^{\Box}$ is the only formula on the succedent somewhere in the derivation below the relevant application of R$\to$ in HML^{-}. (Also, note that L$\to$ is the only rule that lowers the number of formulas on the succedent in a cut-free proof.[17] This is due to the fact that we use only context-sharing rules for $\wedge$ and $\vee$ in modal logics.) Then we can claim the following.

Claim 3.3.4 *If the only case of introducing $A^{\Box} \to B^{\Box}$ on the succedent via $R \to$ are the ones in which $R\Box$ is later applied to $A^{\Box} \to B^{\Box}$ as the only formula on the succedent of a hypersequent occurring lower part of a cut-free proof, then we can replace all of these applications of $R\to$ by $R\to$ to single-conclusion sequents.*

Proof of the claim We prove this claim by induction on the number of the applications of R$\to$ in a given cut-free derivation of a hypersequent whose lowest hypersequent has only formulas in the range of the mapping $()^{\Box}$.

Base case : the number of R$\to$ whose uppersequent and lowersequent have more than one formula on the succedent is 0. In this case, every positive occurrence of a formula of the form $\Box(A^{\Box} \to B^{\Box})$ where $A^{\Box} \to B^{\Box}$ is

---

[17] This is so except the case in which we have more than one occurrence of the same formula on the succedent (e.g. $G|\Gamma A^{\Box} \Rightarrow B^{\Box}, A^{\Box} \to B^{\Box}$ to $G|\Gamma \Rightarrow A^{\Box} \to B^{\Box}$)

introduced by R→ is introduced in the following way.

$$\begin{array}{c}
G^{\Box}|\Gamma', A^{\Box} \Rightarrow B^{\Box} \\ \hline
G^{\Box}|\Gamma' \Rightarrow A^{\Box} \to B^{\Box} \\ \hline\hline
G^{\Box}|\Box\Gamma \Rightarrow A^{\Box} \to B^{\Box} \\ \hline
G^{\Box}|\Box\Gamma \Rightarrow \Box(A^{\Box} \to B^{\Box})
\end{array}\ \mathrm{R}\Box$$

Inductive case: Suppose that we have n applications of R→ in a cut-free proof that has more than one formula on the succedent of the relevant sequent.

Pick up the lowermost application of $R \to$ to a sequent that has more than one formula on the succedent, which looks as follows.[18]

$$\begin{array}{c}
\begin{array}{cc}
\begin{array}{c}
G^{\Box}|\Gamma'', A^{\Box} \Rightarrow B^{\Box}, C^{\Box}, \Delta'' \\ \hline
G^{\Box}|\Gamma'' \Rightarrow A^{\Box} \to B^{\Box}, C^{\Box}, \Delta'' \\ \hline
\text{several steps} \\ \hline
G^{\Box}|\Gamma' \Rightarrow A^{\Box} \to B^{\Box}, C^{\Box}, \Delta'
\end{array}\ \text{(1) R}{\to}{*}
&
G^{\Box}|\Theta', D^{\Box} \Rightarrow
\end{array} \\ \hline
G^{\Box}|\Gamma', \Theta', C^{\Box} \to D^{\Box} \Rightarrow A^{\Box} \to B^{\Box}, \Delta' \quad \text{(2) L}{\to} \\ \hline
\text{several steps} \\ \hline
G^{\Box}|\Box\Gamma, \Box\Delta, \Box\Psi(C^{\Box} \to D^{\Box}) \Rightarrow A^{\Box} \to B^{\Box} \\ \hline
G^{\Box}|\Box\Gamma, \Box\Delta, \Box\Psi(C^{\Box} \to D^{\Box}) \Rightarrow \Box(A^{\Box} \to B^{\Box}) \quad \text{(3) R}\Box{**}
\end{array}$$

We divide the case into two subcases.

Case 1. $A^{\Box} \to B^{\Box}$ does not occur in Δ'' or any other succedent of a sequent on a branch of a proof tree leading to the application of $R\Box^{**}$.

The proof of this case is based on the following observations.

1) It is easy to note that when we have sequential applications of two rules (except $R\Box$) in a cut-free proof whose principal formulas are two different

---

[18] We have a few comments concerning the details. 1. We put Ψ since we can have further applications of some rules there. This does not change the pattern of the argument. 2. Γ' and Δ' must become modalized. Otherwise, the lowest hypesequent may not be in the range of $()^{\Box}$.

formulas, then the proof does not depends on the order of applications of rules. In other words, we can permute applications of these rules. (Including the pair of $R \rightarrow$ and $L \rightarrow$ applied to different formulas.)

2) Also, by the condition that all the formulas derived in the proof must be in the range of the mapping $()^{\Box}$, we have some further conditions on the proof.

2.1) Until (3), we cannot have any other application of $R\Box$ in the proof, since otherwise we would have an application of $R\Box$ when more than one formula are present on the antecedent.

2.2) In order to make a formula $A^{\Box} \rightarrow B^{\Box}$ be in the range of the mapping $()^{\Box}$, we have to apply $R\Box$ to $A^{\Box} \rightarrow B^{\Box}$ on the succedent. This is possible only if we have $A^{\Box} \rightarrow B^{\Box}$ as the only formula on the succedent.[19]

2.3) It is also crucial to note that once $A^{\Box} \rightarrow B^{\Box}$ is introduced, there is no other application of $R \rightarrow$ until the last line.[20] This is because such applications makes it impossible for such a case to be in the range of the mapping $()^{\Box}$. Then we do not have to think about permuting two $R \rightarrow$.

---

[19] This condition is crucial, since if we can keep more than one formula on the succedent or if the formula $A^{\Box} \rightarrow B^{\Box}$ can be moved to the antecedent by L$\rightarrow$, then there is no guarantee that we can always reduce multiple-conclusion $R \rightarrow$'s to single-conclusion $R \rightarrow$'s. In the former case, it is impossible to apply $R\Box$ subsequently. In the latter case, since $A^{\Box} \rightarrow B^{\Box}$ is transformed to another formula that does not have the form $\Box(A^{\Box} \rightarrow B^{\Box})$, we cannot put $A^{\Box} \rightarrow B^{\Box}$ into the range of the mapping $()^{\Box}$.

[20] E.g., consider the following case. An implicational formula $R \rightarrow$ is introduced to the succedent, and after some steps the formula is moved to the antecedent by $L \rightarrow$ applied. Since we have only one formula at the premise of $R\Box$ on the succedent, this is the only case that this may happen. Not only can we make sure that the very formula $A^{\Box} \rightarrow B^{\Box}$ never moves to the antecedent after $A^{\Box} \rightarrow B^{\Box}$ is introduced, but we can also make sure that we do not have any such case in the proof, due to the range constraint of the mapping.

3) By permuting down $R \to$ to $L \to$ (and the other rules), we can reduce the number of the formulas on the succedent when $R \to$ rule is applied as follows.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{G^{\Box}|\Gamma'', A^{\Box} \Rightarrow B^{\Box}, C^{\Box}, \Delta''}{\text{several steps}}}{G^{\Box}|\Gamma', A^{\Box} \Rightarrow B^{\Box}, C^{\Box}, \Delta'} \qquad G^{\Box}|\Theta', D^{\Box} \Rightarrow}{G^{\Box}|\Gamma', \Theta', C^{\Box} \to D^{\Box}, A^{\Box} \Rightarrow B^{\Box}, \Delta'}\,(2)\ \mathrm{L}{\to}}{\text{several steps}}}{G^{\Box}|\Box\Gamma, \Box\Theta, \Box\Psi(C^{\Box} \to D^{\Box}), A^{\Box} \Rightarrow B^{\Box}}}{G^{\Box}|\Box\Gamma, \Box\Theta, \Box\Psi(C^{\Box} \to D^{\Box}) \Rightarrow A^{\Box} \to B^{\Box}}\,(1)\ \mathrm{R}{\to}^{*}}{G^{\Box}|\Box\Gamma, \Box\Theta, \Box\Psi(C^{\Box} \to D^{\Box}) \Rightarrow \Box(A^{\Box} \to B^{\Box})}\,(3)\ \mathrm{R}\Box^{**}$$

Note that R→* is now applied to a sequent whose succedent has only one formula. Hence, permutations of applications guarantee that this case can be reduced to a single-conclusion $R \to$, which is intuitionistically valid.

This sketch of the proof can be made precise by using the induction on the number of hypersequents (lines) between (1) R→* and (3) R□**.

This argument can be made precise by using an inductive argument based on the sum of the number of hypersequents starting from the introduction of $A^{\Box} \to B^{\Box}$ to the succedent of a sequent to the relevant application of R$\Box^{**}$. Since any application of a rule immediately below the topmost occurrence of $A^{\Box} \to B^{\Box}$ in one of the branches can be permuted with R→, we can apply induction hypothesis and show that all the relevant applications of $R \to$ can be reduced to the one immediately above the one of R$\Box^{**}$.

Case 2. $A^{\Box} \to B^{\Box}$ occurs in Δ'' or the succedent of other premise(s) leading to the application of $R\Box^{**}$.

In these cases, it may be possible that the premise has $A^{\Box} \Rightarrow B^{\Box}, A^{\Box} \to B^{\Box}$ and the conclusion has only $A^{\Box} \to B^{\Box}$ on the succedent after applying R$\to$. These cases violate the condition that the premise of R$\to$ has only one formula on the antecedent.

To accommodate these cases, we consider all the chains of sequents that contain $A^{\Box} \to B^{\Box}$ on the succedent whose conclusion lead to the application of R$\Box^{**}$, i.e. the part of the branches of a proof tree containing occurrences of the formula on the succedent.

We apply the same permutation argument that we gave above to all of these occurrences of $A^{\Box} \to B^{\Box}$. We change all the relevant hypersequents of the form $G^{\Box}|\Gamma \Rightarrow A^{\Box} \to B^{\Box}, \Delta$ to $G^{\Box}|\Gamma, A^{\Box} \Rightarrow B^{\Box}, \Delta$ so that we can reduce applications of R$\to$ to the only one immediately above R$\Box^{**}$. Such permutations are possible since in the relevant parts of the branches of the the original proof tree we can never apply R$\Box$ or R$\to$ (other than the one introducing $A^{\Box} \to B^{\Box}$) for the same reason as given in case 1, so the order of applications of rules does not matter.

This argument can be made precise by using an inductive argument based on the sum of the number of hypersequents starting from the introduction of $A^{\Box} \to B^{\Box}$ to the succedent of a sequent to the relevant application of R$\Box^{**}$, which is similar to Case 1. This guarantees that all the relevant applications of $R \to$ can be reduced to the one immediately above R$\Box^{**}$.

By this argument, in any case, we can reduce the number of application of R→ where the premise has more than one formula on the succedent to strictly smaller than n (at least one, but possible smaller).

By applying IH, we can show that all the applications of R→ whose premise has more than one formula on the succedent can be eliminated from the given cut-free proof. (Note that the argument is essentially the same kind as Maehara [100] and Došen [50].) ⊠(claim)

By this claim, we can make sure that although we have multiple-conclusion R→ in the entire systems of modal logics, as far as the derivations of formulas that are in the range of the mapping $()^{\Box}$ are concerned, we can dispense with the applications of R→ to sequents that have more than one formula on the succedent.

Therefore, if we strip off all the modal operators from an entire proof figure of a hypersequent whose formulas are all in the range of $()^{\Box}$, then every application of a rule in the proof figure is a valid inference (possibly with some redundant steps) in each of the superintuitionistic (or intuitionistic) logics at issue here. ⊠ (theorem)

3.4 First-order extensions of modal logics including S4

Among the logics that we dealt with, we consider first-order extensions of the modal logics that includes S4, namely S4, S4.2, S4.3, S5. Let us first give the grammar of the language of first-order modal logic. Let us start from the ordinary first-order language (without function symbols) with n-ary predicate (for any finite n), first-order quantifiers ($\forall$, $\exists$), individual constants c, and countably many individual variables x. (We do not have function symbols in the language.)

$$A := P(\vec{c}, \vec{x})|\bot|\neg A|A_1 \to A_2|A_1 \wedge A_2|A_1 \vee A_2|\Box A|\forall x A|\exists x A$$

1. Hilbert-style axiomatizations can be given by adding the following axioms and rules quantifications to propositional systems. ("$\equiv$" is a syntactic identity.)

Axioms : $\forall x A \to A[x/t]$; $A[x/t] \to \exists x A$;
$\forall x(B \to A) \to (B \to \forall y A[x/y]$ $(x \notin FV(B)$, $y \equiv x$ or $y \notin FV(A))$
$\forall x(A \to B) \to (\exists y A[x/y] \to B)$ $(x \notin FV(B)$, $y \equiv x$ or $y \notin FV(A))$

Rule : (Generalization) $\vdash A$ implies $\vdash \forall y A[x/y]$ $(y \equiv x$ or $y \notin FV(A))$

For each L$\in$ {S4, S4.2, S4.3, S5}, we use a notation ML$\forall$ to stand for the Hilbert-style axiomatization of the first-order version of L.

2. Hypersequent calculi : for each case of propositional hypersequent calculus, we add the following quantifier rules to it.

$$L\exists \quad \frac{G|\Gamma, \varphi(a) \Rightarrow \Delta}{G|\Gamma, \exists x\varphi(x) \Rightarrow \Delta} \qquad R\exists \quad \frac{G|\Gamma \Rightarrow \varphi(t), \Delta}{G|\Gamma \Rightarrow \exists x\varphi(x), \Delta}$$

$$L\forall \quad \frac{G|\Gamma, \varphi(t) \Rightarrow \Delta}{G|\Gamma, \forall x\varphi(x) \Rightarrow \Delta} \qquad R\forall \quad \frac{G|\Gamma \Rightarrow \varphi(a), \Delta}{G|\Gamma \Rightarrow \forall x\varphi(x), \Delta}$$

For $L\exists$ and $R\forall$, we have the ordinary eigenvariable condition.

We call these system generically HML∀. We first state deductive equivalence between Hilbert-style systems and hypersequent systems (a proof is straightforward). We can naturally extend our interpretation function $\mathcal{I}$ to the first-order case.

Proposition 3.4.1 *ML*∀ $\vdash \Box(\bigwedge\Gamma_1 \to \bigvee\Delta_1) \vee \cdots \vee \Box(\bigwedge\Gamma_n \to \bigvee\Delta_n)$ *iff* *HML*∀ $\vdash \Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_n \Rightarrow \Delta_n$.

Note: We have a few remarks on the first-order extensions of these logics.

1. It turns out that these logics, in particular S4.2 and S4.3 cases, are very rare cases that are complete with respect to a natural class of Kripke models that are natural extensions of propositional versions of these logics. Ghilardi [66] showed that for any any $L\forall \supsetneq S5\forall$ and $L\forall \supsetneq S4.3\forall$ are incomplete w.r.t. Kripke semantics. This means that essentially all simple cases of first-order modal logics between S4 and S5 (Kripke complete cases) are handled here, since it is probably difficult to find out simple formulations of first-order modal logics L s.t. $S4\forall \subsetneq L \subsetneq S4.2\forall$.

2. Since these are extensions of S4, under the current formulations of free variables, converse Barcan formulas are derivable.[21] Only for S5 case, Barcan formulas are derivable as follows (in cut-free HS5$\forall$). Note that Barcan formulas are not derivable in S4.3$\forall$.[22]

$$\begin{array}{c}
A(x) \Rightarrow A(x) \\ \hline
\Box A(x) \Rightarrow A(x) \\ \hline
\Box A(x) \Rightarrow \mid \Rightarrow A(x) \\ \hline
\forall x \Box A(x) \Rightarrow \mid \Rightarrow A(x) \\ \hline
\forall x \Box A(x) \Rightarrow \mid \Rightarrow \forall x A(x) \\ \hline
\forall x \Box A(x) \Rightarrow \mid \Rightarrow \Box \forall x A(x) \\ \hline
\forall x \Box A(x) \Rightarrow \Box \forall x A(x) \mid \Rightarrow \Box \forall x A(x) \\ \hline
\forall x \Box A(x) \Rightarrow \Box \forall x A(x) \mid \forall x \Box A(x) \Rightarrow \Box \forall x A(x) \\ \hline
\forall x \Box A(x) \Rightarrow \Box \forall x A(x)
\end{array}$$

Let us now move on to the issue of cut-elimination. In first-order cases, the substitutivity of the quantifier rules does not hold in the same as it does

---

[21] It is possible to formulate a system in which converse Barcan is not derivable, but then the proof systems may look unnatural [95]. Also, if we keep the current formulation, it is more convenient to discuss extensions of Gödel embedding to first-order logic.

[22] This issue of Barcan formulas brings in some subtlety when it comes to Gödel embedding. It is not the case that our syntactic proofs of soundness and faithfulness of Gödel translations of these logics can be extended to first-order cases. For S4 and S4.2, the proof can be extended to first-order case since both intuitionistic logic and logic of weak excluded middle semantically have monotonic domain, which coincides with the case of modal logics. For S5, the proof can be extended since there is no way that classical logic invalidates formula that corresponds to constant domain condition $\forall x(A \vee B(x)) \to (A \vee \forall x B(x))$ (where $x \notin FV(A)$). However, for S4.3 and Gödel-Dummett logic, the extension breaks down since logic of S4.3 does not enforce the derivability of Barcan formulas, but the hypersequent calculus for Gödel-Dummett logic derives constant domain formula. (First-order Gödel-Dummett logic is usually axiomatized by using constant domain formulas, and the hypersequent calculus for Gödel-Dummett logic naturally derives constant domain formulas. However, it is possible to formulate a similar logic without using constant domain formulas.) If one does not care about syntactic proofs in decent cut-free proof systems, it may be easy to fix the discrepancy since we can just add Barcan formulas to the Hilbert-style system for S4.3 (the system must be still complete with respect to suitable constant domain Kripke semantics). Then the correspondence is restored, but proof-theoretically it is not clear how we can add a natural rule corresponding to Barcan formulas in the hypersequent formulation of S4.3.

for propositional cases, due to the eigenvariable condition. Also, formulas $\varphi(a)$ and $\varphi(t)$ occurring in the premises may not be identical; hence, it is not directly possible to apply IH of cut to these formulas whose complexity is smaller than those with quantifiers.

However, the problem can be avoided by proving the following lemma. (Here the notation $d; G_1(a), \ldots, G_1(a) \vdash_{\mathsf{HML}\forall} G(a)$ means that d is a derivation of a hypersequent $G(a)$ having $G_1(a), \ldots, G_n(a)$ in each step.

Lemma 3.4.2 *If $d; G_1(a), \ldots, G_n(a) \vdash_{\mathit{HML}\forall} G(a)$ and t is a term with variables not occurring in d, then $d'; G_1(t), \ldots, G_n(t) \vdash_{\mathit{HML}\forall} G(t)$ with length of d = length of d'.*

Proof Proof is by simple induction on the length of derivations.

With this lemma, substitutivity of the quantifier rule can hold.

We now prove the quantifier cases of lemma 3.2.2 and lemma 3.2.3

Proof (proof of lemma 3.2.2) Case 8. The last inference of d_l is $L\forall$. Let A be $\forall x\varphi(x)$.

$$L\forall\ \frac{G|\Gamma_1, \varphi(t), [A]^{\lambda_1 - 1} \Rightarrow \Delta_1| \ldots |\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}{G|\Gamma_1, \forall x\varphi(x), [A]^{\lambda_1 - 1} \Rightarrow \Delta_1| \ldots |\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}$$

Note that the derivation d_r ends with the following inference

$$\frac{H|\Pi \Rightarrow \varphi(a), \Sigma}{H|\Pi \Rightarrow \forall x\varphi(x), \Sigma}, \text{ where } a \text{ does not occur in } \Pi \text{ or } \Sigma.$$

By IH, $G|H|\Gamma_1, \varphi(t), \Pi^{\lambda_1 - 1} \Rightarrow \Delta_1, \Sigma^{\lambda_1 - 1}| \ldots |\Gamma_n, \Pi^{\lambda_n} \Rightarrow \Sigma^{\lambda_n}, \Delta_n$.

By applying the lemma to $d_r \vdash H|\Pi \Rightarrow \varphi(t), \Sigma$ and by applying cut to this w.r.t. $\varphi(t)$, we derive the following: $G|H|\Gamma_1, \Pi^{\lambda_1} \Rightarrow \Delta_1, \Sigma^{\lambda_1}| \ldots |\Gamma_n, \Pi^{\lambda_n} \Rightarrow \Sigma^{\lambda_n}, \Delta_n$.

This is the desired hypersequent (namely obtained by applying cut to the lowest hypersequents in d_l and d_r), but the cut-rank of the derivation is $\leq |A|$ (which is lower than the cut just mentioned in the parenthesis).

Case 9. The last inference is $L\exists$. $A = \exists x \varphi(x)$.

$$L\exists \ \frac{G|\Gamma_1, \varphi(a), [A]^{\lambda_1 - 1} \Rightarrow \Delta_1| \ldots |\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}{G|\Gamma_1, \exists x \varphi(x), [A]^{\lambda_1 - 1} \Rightarrow \Delta_1| \ldots |\Gamma_n, [A]^{\lambda_n} \Rightarrow \Delta_n}$$

The derivation d_r ends with the following inference

$$\frac{H|\Pi \Rightarrow \varphi(t), \Sigma}{H|\Pi \Rightarrow \exists x \varphi(x), \Sigma}.$$

By IH, $H|G|\Gamma_1, \varphi(a), \Pi^{\lambda_1 - 1} \Rightarrow \Delta_1, \Sigma^{\lambda_1 - 1}| \ldots |\Gamma_n, \Pi^{\lambda_n} \Rightarrow \Sigma^{\lambda_n}, \Delta_n$

Applying the lemma to this hypersequent and replacing $\varphi(a)$ and $\varphi(t)$, we get

$G|H|\Gamma_1, \varphi(t), \Pi^{\lambda_1 - 1} \Rightarrow \Delta_1, \Sigma^{\lambda_1 - 1}| \ldots |\Gamma_n, \Pi^{\lambda_n} \Rightarrow \Sigma^{\lambda_n}, \Delta_n$ Then, apply cut to the two hypersequents with respect to $\varphi(t)$.

We get $G|H|H|\Gamma_1, \varphi(t), \Pi^{\lambda_1 - 1} \Rightarrow \Delta_1, \Sigma^{\lambda_1 - 1}| \ldots |\Gamma_n, \Pi^{\lambda_n} \Rightarrow \Sigma^{\lambda_n}, \Delta_n$.

Applying EC, $G|H|\Gamma_1, \Pi^{\lambda_1} \Rightarrow \Delta_1, \Sigma^{\lambda_1}| \ldots |\Gamma_n, \Pi^{\lambda_n} \Rightarrow \Sigma^{\lambda_n}, \Delta_n$.

This is the desired hypersequent, but the cut rank of the derivation is

$\leq |A|$.

Proof (Proof of lemma 3.2.3) Case 8. d_r ends with $R\exists$. $A = \exists x\varphi(x)$.

$$R\exists\ \frac{G|\Pi_1 \Rightarrow \varphi(t), [A]^{\lambda_1 - 1}, \Sigma_1| \ldots |\Pi_n \Rightarrow [A]^{\lambda_n}, \Sigma_n}{G|\Pi_1 \Rightarrow \exists x\varphi(x), [A]^{\lambda_1 - 1}, \Delta_1| \ldots |\Gamma_n \Rightarrow [A]^{\lambda_n}, \Sigma_n}$$

The derivation d_l ends with the following inference

$$\frac{H|\Gamma, \varphi(a) \Rightarrow \Delta}{H|\Gamma, \exists x\varphi(x) \Rightarrow \Delta}, \text{ where } a \text{ does not occur in } \Pi \text{ or } \Sigma.$$

Applying the lemma to the premise of the last inference of d_l, we get $H|\Gamma, \varphi(t) \Rightarrow \Delta$ by replacing $\varphi(a)$ by $\varphi(t)$.

Applying IH, $G|H|\Pi_1, \Gamma^{\lambda_1 - 1} \Rightarrow \varphi(t), \Sigma_1, \Delta^{\lambda_1 - 1}| \ldots |\Pi_n, \Gamma^{\lambda_n} \Rightarrow \Sigma_n, \Delta^{\lambda_n}$.

By cut with respect to $\varphi(t)$, we get

$G|G|H|\Pi_1, \Gamma^{\lambda_1} \Rightarrow \Sigma_1, \Delta^{\lambda_1}| \ldots |\Pi_n, \Gamma^{\lambda_n} \Rightarrow \Sigma_n, \Delta^{\lambda_n}$.

Using EC, we get the desired hypersequent $G|H|\Pi_1, \Gamma^{\lambda_1} \Rightarrow \Sigma_1, \Delta^{\lambda_1}| \ldots |\Pi_n, \Gamma^{\lambda_n} \Rightarrow \Sigma_n, \Delta^{\lambda_n}$. Note that the derivation of this has cut rank only with $\leq |A|$.

Case 9. d_r ends with $R\forall$. $A = \forall x\varphi(x)$.

$$R\forall\ \frac{G|\Pi_1 \Rightarrow \varphi(a), [A]^{\lambda_1 - 1}, \Sigma_1| \ldots |\Pi_n \Rightarrow [A]^{\lambda_n}, \Sigma_n}{G|\Pi_1 \Rightarrow \forall x\varphi(x), [A]^{\lambda_1 - 1}, \Delta_1| \ldots |\Gamma_n \Rightarrow [A]^{\lambda_n}, \Sigma_n}$$

The derivation d_l ends with the following inference

$$\frac{H|\Gamma, \varphi(t) \Rightarrow \Delta}{H|\Gamma, \forall x\varphi(x) \Rightarrow \Delta}$$

We apply the lemma to the premise of d_r, replacing $\varphi(a)$ and $\varphi(t)$, and then we get $G|\Pi_1 \Rightarrow \varphi(t), [A]^{\lambda_1 - 1}, \Sigma_1| \dots |\Pi_n \Rightarrow [A]^{\lambda_n}, \Sigma_n$.

Applying IH, we get $G|H|\Pi_1, \Gamma^{\lambda_1 - 1} \Rightarrow \varphi(t), \Delta^{\lambda_1 - 1}, \Sigma_1| \dots |\Pi_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Sigma_n$.

Using cut with respect to $\varphi(t)$, we get the hypersequent $G|G|H|\Pi_1, \Gamma^{\lambda_1} \Rightarrow \Delta^{\lambda_1}, \Sigma_1| \dots |\Pi_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Sigma_n$. By EC, we get $G|H|\Pi_1, \Gamma^{\lambda_1} \Rightarrow \Delta^{\lambda_1}, \Sigma_1| \dots |\Pi_n, \Gamma^{\lambda_n} \Rightarrow \Delta^{\lambda_n}, \Sigma_n$. This is the desired hypersequent, and the cut rank of the derivation is $\leq |A|$. ⊠

Adding these cases to the lemma 3.2.3. leads to the extension of cut-elimination to the first-order cases.

Theorem 3.4.3 (Cut-elimination) *For all first-order extension of HML∀ where L∈ {S4,S4.2,S4.3,S5}, cut-elimination holds for HML∀.*

Proof Proof of cut-elimination is similar to the propositional cases. ⊠

Chapter 4

Tableaux and Hypersequents for Justification Logics

4.1 Tableaux and Hypersequents for S4LPN

The first system of Justification Logic, the Logic of Proofs (LP), is introduced by Artemov ([6]) as a logic that can explicitly talk about proofs. An earlier sketch of the Logic of Proofs was suggested by Gödel in [70]. Several variants have been studied in combination with traditional modal logics. One such variant is S4LP, which was introduced by Artemov and Nogina ([8]) and also studied by Fitting ([58]). This logic contains knowledge modality $\Box F$ and justification assertions $t : F$. Other examples are LPP ([178]) and GLA ([109]), both of which are combination of LP and provability logic (GL, or

Gödel-Löb logic). Artemov and Nogina in [8][1] introduced both logics, S4LP and S4LPN, using Hilbert-style axiomatic systems. The latter is S4LP with mixed negative introspection $\neg t\!:\!F \rightarrow \Box\neg t\!:\!F$. Fitting[58] and Renne[135] found destructive tableau systems for S4LP. But so far no tableau system or sequent calculus for S4LPN or GLA has been proposed. Moreover, since such a formula as $\Box t\!:\!F \vee \Box\neg t\!:\!F$ is a theorem of S4LPN and GLA, the task of finding cut-free destructive tableau systems for these logics seems to be hopeless. Here we suggest more flexible frameworks to give cut-free complete proof-systems for S4LPN. We first give a prefixed tableau system for S4LPN in the sense of [57]. However, from a philosophical perspective, one might think that a prefixed tableau system contain too much semantic information in the form of prefixes (cf., Avron [18]). To overcome this potential weakness, we formulate a hypersequent calculus for the logic (HS4LPN) and show that there is a way of translating a closed prefixed tableau to a proof in HS4LPN. The translation is done via the following steps. First, we convert a given prefixed tableau proof to a certain normal form. Second, we translate the proof in normal form into a proof in an auxiliary tableau system. Then we translate the tableau proof in the auxiliary system to a (cut-free) proof in the HS4LPN. This gives us a version of the completeness theorem for HS4LPN without cut. As a corollary, we will obtain a semantic proof of cut-admissibility for HS4LPN.

In the final subsection of this section, we discuss a subsystem of the pre-

---

[1] We mostly follow the notation and the terminology of [8] concerning S4LPN.

fixed tableau for S4LPN, i.e., S4LP, since it turns out that we can formulate a hypersequent calculus for the logic (HS4LP) and apply a similar method of translation from the prefixed tableau system to HS4LP, which is simpler than the one used for the the systems for S4LPN.

4.1.1 A Hilbert-style system, a prefixed tableau system, and semantics for **S4LPN**

We give the language of S4LPN and a Hilbert-style axiom system for S4LPN.

The language of S4LPN is specified as follows:

1. The class of proof terms in the language of S4LPN is specified as follows.

$$t := x|a|!t|t_1 \cdot t_2|t_1 + t_2.$$

2. The class of formulas in the language of S4LPN is specified as follows:

$$A := P_i|\bot|\neg A|A_1 \to A_2|A_1 \wedge A_2|A_1 \vee A_2|t{:}A|\Box A$$

In this section, "Trm" stands for the set of proof-terms in the language of S4LPN. "Fmla" stands for the set of formulas in the language of S4LPN.

A *constant specification* is a mapping $\mathcal{CS}$ from proof constants to sets of formulas (possibly empty). A formula A *has a proof constant* c with respect

$\mathcal{CS}$ if $A \in \mathcal{CS}(c)$.

A Hilbert style system of S4LPN is as follows.

Axioms 0) Axioms of Propositional Logic

1) Axioms of explicit knowledge:

1. $t\!:\!\varphi \to \varphi$
2. $t\!:\!(\varphi \to \psi) \to (s\!:\!\varphi \to t \cdot s\!:\!\psi)$
3. $t\!:\!\varphi \to !t\!:\!t\!:\!\varphi$
4. $t\!:\!\varphi \to t + s\!:\!\varphi$, $s\!:\!\varphi \to t + s\!:\!\varphi$

2) Axioms of implicit knowledge:

1. $\Box\varphi \to \varphi$
2. $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$
3. $\Box\varphi \to \Box\Box\varphi$

3) Connecting Axiom:

1. $t\!:\!\varphi \to \Box\varphi$
2. $\neg t\!:\!\varphi \to \Box\neg t\!:\!\varphi$

Rules of Inference

1. Modus Ponens
2. Necessitation $\varphi/\Box\varphi$

3. $c\!:\!A$ (Axiom Necessitation), where A is one of the above axioms and $A \in \mathcal{CS}(c)$.

Now we present the prefixed tableau system (We call this TS4LPN). We follow Fitting's terminology for basic notions in the prefixed tableau system.[2] In particular, our prefix is a finite sequence of positive integers that has only 1 for the initial element of a sequence of natural numbers. In $\sigma T\varphi$ or $\sigma F\varphi$, σ is a prefix of a signed formula $T\varphi$ or $F\varphi$. $\sigma.n$ is called a simple extension of σ. σ' is accessible from σ iff $\sigma \leq \sigma'$ ($\leq$ means that "is an (not necessarily proper) initial segment of"). σ' is e-accessible from σ iff $1 \leq \sigma$ and $1 \leq \sigma'$. (Note that such e-accessibility is an equivalence relation.) Also, we use the following terminology for "used" for a technical reason related to our proof of the completeness theorem.[3] A prefix σ is *used* on a branch if a prefix of which it is an initial segment has already occurred on the branch. Otherwise, a prefix is *new* on a branch. We say φ has a prefixed tableau proof if there is a closed prefixed tableau starting from $1F\varphi$.

Rules for Classical Propositional Logic.

α-rule:

$$\frac{\sigma T\varphi \wedge \psi}{\begin{array}{c}\sigma T\varphi\\ \sigma T\psi\end{array}} \qquad \frac{\sigma F\varphi \vee \psi}{\begin{array}{c}\sigma F\varphi\\ \sigma F\psi\end{array}} \qquad \frac{\sigma F\varphi \rightarrow \psi}{\begin{array}{c}\sigma T\varphi\\ \sigma F\psi\end{array}}$$

---

[2] The origin of the prefixed tableau system goes back to Fitting ([57]), but the modal part of the system adopted here comes from [59] and it is based on Massacci's single step tableaux system ([102]). See, Goré ([75]), too.

[3] The use was suggested by Prof. Fitting in personal communication. This significantly made simpler the current proof of completeness. Note that for any ordinary construction of a tableau, this does not make any difference from more traditional terminology of "used" since we always construct a prefix from 1 in a step by step manner.

$$\frac{\sigma T \neg\varphi}{\sigma F \varphi} \qquad \frac{\sigma F \neg\varphi}{\sigma T \varphi}$$

β-rule:

$$\frac{\sigma T \varphi \vee \psi}{\sigma T \varphi \mid \sigma T \psi} \qquad \frac{\sigma F \varphi \wedge \psi}{\sigma F \varphi \mid \sigma F \psi} \qquad \frac{\sigma T \varphi \to \psi}{\sigma F \varphi \mid \sigma T \psi}$$

$\bot$-rule: since we have $\bot$ in the language, we formulate the following rule for $\bot$.

A branch is closed if $\sigma T \bot$ occurs on it.

Rules for LP: (Explicit S4 ν-rules)

EK $\dfrac{\sigma T t{:}\varphi}{\sigma.n T \varphi}$ ($\sigma.n$ is used.) ET $\dfrac{\sigma T t{:}\varphi}{\sigma T \varphi}$

E4 $\dfrac{\sigma T t{:}\varphi}{\sigma.n T t{:}\varphi}$ ($\sigma.n$ is used.) E4r $\dfrac{\sigma.n T t{:}\varphi}{\sigma T t{:}\varphi}$

EF $\dfrac{\sigma F t{:}\varphi}{\sigma.n F t{:}\varphi}$ ($\sigma.n$ is used.) EFr $\dfrac{\sigma.n F t{:}\varphi}{\sigma F t{:}\varphi}$

Operational Rules on F's :

!-rule $\dfrac{\sigma F !t{:}t{:}\varphi}{\sigma F t{:}\varphi}$ $\cdot$-rule $\dfrac{\sigma F (s\cdot t){:}\varphi}{\sigma F t{:}\psi\to\varphi | \sigma F s{:}\psi}$

+-rule $\dfrac{\sigma F (s+t){:}\varphi}{\sigma F t{:}\varphi}$ $\dfrac{\sigma F (t+s){:}\varphi}{\sigma F t{:}\varphi}$

Modal Rules:

ν-rules: $\nu_K\ \dfrac{\sigma T\Box\varphi}{\sigma.n T\varphi}$ ($\sigma.n$ is used.) $\nu_T\ \dfrac{\sigma T\Box\varphi}{\sigma T\varphi}$

$\nu_4\ \dfrac{\sigma T\Box\varphi}{\sigma.n T\Box\varphi}$ ($\sigma.n$ is used.)

π-rule: $\dfrac{\sigma F\Box\varphi}{\sigma.n F\varphi}$ ($\sigma.n$ is new.)

In addition, we have Constant Specification Rules as follows: a branch is closed if it has $\sigma F c{:}A$, where A is an axiom of $\mathsf{S4LPN}$ and $A \in \mathcal{CS}(c)$.

Remark: The rules EF and EFr, which we call "reverse rules" in the following, may need some discussion. These are indeed sound with respect to Kripke semantics that we define later. On the other hand, with the rule of cut, i.e.,

Cut $\dfrac{}{\sigma T\varphi \mid \sigma F\varphi}$, the prefixed tableau system is complete with respect to the appropriate Kripke semantics without those rules. This may give an impression that these rules are redundant. However, it seems that without those rules, we cannot avoid using cut at some point in a proof of completeness.

Next, we define Fitting-style Kripke semantics for S4LPN. Let a triple (K, R, R^e) be a frame, where K is non-empty set, R is a reflexive and transitive relation on K, and R^e is a reflexive, symmetric and transitive relation on K. We also assume that $R \subseteq R^e$.

Let $\mathcal{E}$ be an evidence function: $Trm \longrightarrow \mathcal{P}(Fmla)$ (with respect to a constant specification $\mathcal{CS}$) that satisfies the following properties:

1. uR^ev implies $\mathcal{E}(u,t) \subseteq \mathcal{E}(v,t)$ (Monotonicity)[4]
2. $F \to G \in \mathcal{E}(u,t)$ and $F \in \mathcal{E}(u,s)$ implies $G \in \mathcal{E}(u, t \cdot s)$
3. $F \in \mathcal{E}(u,t)$ implies $t{:}F \in \mathcal{E}(u, !t)$
4. $\mathcal{E}(u,s) \cup \mathcal{E}(u,t) \subseteq \mathcal{E}(u, s+t)$
5. $\mathcal{CS}(c) \subseteq \mathcal{E}(u,c)$

Then, a Kripke model $\mathcal{K}$ (with respect to a constant specification $\mathcal{CS}$) can be defined as a quintuple $(K, R, R^e, \mathcal{E}, \mathcal{V})$. $\mathcal{V}$ is a function from propositional variables to subsets of K. We also define a forcing $\Vdash$ as a relation on $K \times Fmla$ that satisfies the following inductive property.

---

[4] Symmetry of uR^ev actually implies $\mathcal{E}(u,t) = \mathcal{E}(v,t)$.

0. $u \Vdash p$ if and only if $u \in \mathcal{V}(p)$; $u \nVdash \bot$ for all $u \in K$.
1. $\Vdash$ commutes with propositional connectives $\{\wedge, \vee, \rightarrow, \neg\}$ at each state.
2. $u \Vdash \Box\varphi$ iff for every $v \in K$, s.t. uRv, $v \Vdash \varphi$
3. $u \Vdash t{:}\varphi$ iff $\varphi \in \mathcal{E}(u,t)$ and for every $v \in K$, s.t. $uR^e v$, $v \Vdash \varphi$.
4. $A \in \mathcal{CS}(c)$ implies $\mathcal{K}, u \Vdash c{:}A$ for every $u \in K$.

We have the following theorem from Artemov and Nogina [8].

Theorem 4.1.1 *S4LPN is sound and complete with respect to the Kripke semantics defined above.*

Here we introduce some terminology about a connection between the language used in tableaux and Kripke semantics. A signed formula $F\varphi$, $T\varphi$ (written Φ schematically) is *realized* at a possible world u of a model $\mathcal{K}$ if 1) the formula is $T\varphi$ and $\mathcal{K}, u \Vdash \varphi$ or 2) the formula is $F\varphi$ and $\mathcal{K}, u \nVdash \varphi$. A set S of prefixed, signed formulas is *satisfiable* if there is a model $\mathcal{K}$ and a mapping $\mathcal{N}$ (called "interpretation") from the prefixes in S to possible worlds in $\mathcal{K}$, such that if $\sigma\Phi \in S$, then Φ is realized at $\mathcal{N}(\sigma)$ in $\mathcal{K}$, where Φ is a signed formula and such an $\mathcal{N}$ satisfies the condition (1) $\sigma \leq \sigma' \Longrightarrow \mathcal{N}(\sigma)R\mathcal{N}(\sigma')$ and (2) $1 \leq \sigma$ and $1 \leq \sigma' \Longrightarrow \mathcal{N}(\sigma)R^e\mathcal{N}(\sigma')$. A tableau branch is *satisfiable* if the set of prefixed formulas on it is satisfiable. A tableau is *satisfiable* if some tableau branch is.

4.1.2 Soundness and Completeness of the Prefixed Tableau System

We prove soundness and completeness of the prefixed tableau system.

Lemma 4.1.2 *Suppose $\mathcal{T}$ is a satisfiable tableau. If any tableau rule for $\mathsf{S4LPN}$ is applied to $\mathcal{T}$, then the resulting tableau is still satisfiable.*

Proof Suppose a tableau is $\mathsf{S4LPN}$-satisfiable because a branch θ of $\mathcal{T}$ is $\mathsf{S4LPN}$-satisfiable, i.e., its members are realized at $\mathcal{N}(\sigma)$ of model $\mathcal{K}$. Suppose that a tableau rule for $\mathsf{S4LPN}$ is applied to the tableau $\mathcal{T}$. (We call the resulting tableau $\mathcal{T}'$.) The entire proof is divided into two cases:

Case 1: Our tableau rule is not applied on the branch θ. Then, θ is still present on the new tableau and θ is satisfiable, which makes $\mathcal{T}'$ obviously satisfiable.

Case 2: Our tableau rule is applied on the branch θ. Here we treat only some explicit cases (since $\Box$ cases are standard).

Rules for LP (Explicit ν-rules):[5]

EK : Suppose $\sigma Tt:\varphi$ occurs on θ, the rule EK is applied and $\sigma.nT\varphi$ ($\sigma.n$ is used) is added on θ. By the assumption of satisfiability of θ, $Tt:\varphi$ is realized at $\mathcal{N}(\sigma)$ of a model $\mathcal{K}$. So, $\mathcal{N}(\sigma) \Vdash t:\varphi$.[6] On the other hand,

---

[5] The cases of E4r and EF essentially use symmetry of R^e. For the other cases of E rules, $\sigma \leq \sigma.n$ or $\sigma \leq \sigma$ is sufficient.

[6] In the following, we omit $\mathcal{K}$ unless we have a case where the omission can cause some confusion.

our sequence $\sigma.n$ must have been used before, so $\mathcal{N}$ for $\sigma.n$ must be already defined and since $1 \leq \sigma$ and $1 \leq \sigma.n$, $\mathcal{N}(\sigma)R^e\mathcal{N}(\sigma.n)$ By the truth condition of an explicit modal formula, we have $\mathcal{N}(\sigma) \Vdash t : \varphi$ iff $\varphi \in \mathcal{E}(\mathcal{N}(\sigma), t)$ and $\forall v(\mathcal{N}(\sigma)R^e v \Rightarrow v \Vdash \varphi)$. So, $\mathcal{N}(\sigma)R^e\mathcal{N}(\sigma.n)$ implies $\mathcal{N}(\sigma.n) \Vdash \varphi$. Hence $\mathcal{N}(\sigma.n) \Vdash \varphi$. Then, $T\varphi$ is realized at $\mathcal{N}(\sigma.n)$. So, $\theta \cup \{\sigma.nT\varphi\}$ is satisfiable.

ET : Suppose $\sigma Tt : \varphi$ occurs in θ, ET rule is applied and $\sigma T\varphi$ is added on θ. By the assumption of satisfiability of θ, $Tt : \varphi$ is realized at $\mathcal{N}(\sigma)$ in a model $\mathcal{K}$. So, $\mathcal{N}(\sigma) \Vdash t : \varphi$. By the truth condition of $t : \varphi$, $\varphi \in \mathcal{E}(\mathcal{N}(\sigma), t)$ and for all v, s.t. $\mathcal{N}(\sigma)R^e v$, $v \Vdash \varphi$. So, $\mathcal{N}(\sigma)R^e\mathcal{N}(\sigma) \Rightarrow \mathcal{N}(\sigma) \Vdash \varphi$. On the other hand, since obviously $1 \leq \sigma$ and $1 \leq \sigma$, $\mathcal{N}(\sigma)R^e\mathcal{N}(\sigma)$. So, $T\varphi$ is realized at $\mathcal{N}(\sigma)$ in $\mathcal{K}$. Hence, $\theta \cup \{\sigma T\varphi\}$ is satisfiable.

E4 : Suppose $\sigma Tt : \varphi$ occurs on θ, E4 is applied and $\sigma.nTt : \varphi$ is added on θ. (Here $\sigma.n$ is used.) By the assumption of the satisfiability of θ, $Tt : \varphi$ is realized at $\mathcal{N}(\sigma)$ in $\mathcal{K}$. So, $\mathcal{N}(\sigma) \Vdash t : \varphi$. Since $1 \leq \sigma$ and $1 \leq \sigma.n$, $\mathcal{N}(\sigma)R^e\mathcal{N}(\sigma.n)$. [Since $\sigma.n$ is used, $\mathcal{N}$ is already defined for $\sigma.n$.] By the truth condition of $t : \varphi$, $\varphi \in \mathcal{E}(\mathcal{N}(\sigma), t)$ and for all v, s.t. $\mathcal{N}(\sigma)R^e v$, $v \Vdash \varphi$. Now we want to show that $\varphi \in \mathcal{E}(\mathcal{N}(\sigma.n), t)$ and for all v, s.t. $\mathcal{N}(\sigma.n)R^e v$, $v \Vdash \varphi$. By monotonicity, $\varphi \in \mathcal{E}(\mathcal{N}(\sigma), t)$ and $\mathcal{N}(\sigma)R^e\mathcal{N}(\sigma.n)$ implies $\varphi \in \mathcal{E}(\mathcal{N}(\sigma.n), t)$. So, $\varphi \in \mathcal{E}(\mathcal{N}(\sigma.n), t)$ holds. Then, suppose $\mathcal{N}(\sigma.n)R^e v$. Since $\mathcal{N}(\sigma)R^e\mathcal{N}(\sigma.n)$, $\mathcal{N}(\sigma)R^e v$ by transitivity. $\mathcal{N}(\sigma)R^e v$ implies $v \Vdash \varphi$, due to the truth condition of $t : \varphi$ at $\mathcal{N}(\sigma)$. So, $v \Vdash \varphi$. Hence, for any v, s.t. $\mathcal{N}(\sigma.n)R^e v$, $v \Vdash \varphi$. Therefore, $\mathcal{N}(\sigma.n) \Vdash t : \varphi$. So, $Tt : \varphi$ is realized at $\mathcal{N}(\sigma.n)$ in $\mathcal{K}$. So, $\theta \cup \{\sigma.nTt : \varphi\}$ is satisfiable.

E4r : Suppose $\sigma.nTt{:}\varphi$ occurs on θ, E4r is applied and $\sigma Tt{:}\varphi$ is added on θ. By the assumption of the satisfiability of θ, $Tt{:}\varphi$ is realized at $\mathcal{N}(\sigma.n)$ in $\mathcal{K}$. So, $\mathcal{N}(\sigma.n) \Vdash t : \varphi$. Since $1 \leq \sigma$ and $1 \leq \sigma.n$, $\mathcal{N}(\sigma)R^e\mathcal{N}(\sigma.n)$ (as $\sigma.n$ is used and $\mathcal{N}$ is already defined.) Also, by symmetry of R^e, we have $\mathcal{N}(\sigma.n)R^e\mathcal{N}(\sigma)$.

By the truth condition of $t : \varphi$, $\varphi \in \mathcal{E}(\mathcal{N}(\sigma.n), t)$ and for any v, s.t. $\mathcal{N}(\sigma.n)R^ev$, $v \Vdash \varphi$. By monotonicity, $\varphi \in \mathcal{E}(\mathcal{N}(\sigma), t)$. The first part is done.

Suppose $\mathcal{N}(\sigma)R^ev$. Since $\mathcal{N}(\sigma.n)R^e\mathcal{N}(\sigma)$, by transitivity, $\mathcal{N}(\sigma.n)R^ev$. So, by the second part of the above truth condition, we have $v \Vdash \varphi$. So, for any v, s.t. $\mathcal{N}(\sigma)R^ev$, $v \Vdash \varphi$. The second part is done. Hence, $\mathcal{N}(\sigma) \Vdash t{:}\varphi$. So, $Tt{:}\varphi$ is realized at $\mathcal{N}(\sigma)$ in $\mathcal{K}$. Therefore, $\theta \cup \{\sigma Tt{:}\varphi\}$ is satisfiable.

EF : Suppose $\sigma Ft{:}\varphi$ occurs on θ, EF is applied and $\sigma.nFt{:}\varphi$ is added on θ. (Here $\sigma.n$ is used.) By the assumption of the satisfiability of θ, $Ft{:}\varphi$ is realized at $\mathcal{N}(\sigma)$ in $\mathcal{K}$. So, $\mathcal{N}(\sigma) \nVdash t : \varphi$. Since clearly $1 \leq \sigma$ and $1 \leq \sigma.n$, $\mathcal{N}(\sigma)R^e\mathcal{N}(\sigma.n)$. [Since $\sigma.n$ is used, $\mathcal{N}$ is already defined for $\sigma.n$.] By symmetry of R^e, $\mathcal{N}(\sigma.n)R^e\mathcal{N}(\sigma)$. By the truth condition of $t : \varphi$, (1) $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$ or (2) for some v, s.t. $\mathcal{N}(\sigma)R^ev$ and $v \nVdash \varphi$. Now we want to show that $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma.n), t)$ or for some v, s.t. $\mathcal{N}(\sigma.n)R^ev$ and $v \nVdash \varphi$. By (1) $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$, monotonicity and $\mathcal{N}(\sigma.n)R^e\mathcal{N}(\sigma)$ imply $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma.n), t)$. So, we can derive $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma.n), t)$ or for some v, s.t. $\mathcal{N}(\sigma.n)R^ev$ and $v \nVdash \varphi$. We derive the second part from (2) now. Pick a world that satisfies the condition and temporarily call it v_1. Then, $\mathcal{N}(\sigma)R^ev_1$ and $v_1 \nVdash \varphi$. By transitivity, $\mathcal{N}(\sigma.n)R^ev_1$. Hence, for some v, s.t. $\mathcal{N}(\sigma.n)R^ev$ and $v \nVdash \varphi$. So, again, $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma.n), t)$ or for some v, s.t. $\mathcal{N}(\sigma.n)R^e v$ and $v \nVdash \varphi$. Hence, $\mathcal{N}(\sigma.n) \nVdash t\!:\!\varphi$. So, $Ft\!:\!\varphi$ is realized at $\mathcal{N}(\sigma.n)$ in $\mathcal{K}$. Therefore, $\theta \cup \{\sigma.nFt\!:\!\varphi\}$ is satisfiable.

Operational Rules:

!-rule: Suppose $\sigma F!t : t : \varphi$ occurs in θ, !-rule is applied on $!t : t : \varphi$ and $\sigma Ft : \varphi$ is added on θ. By the assumption of satisfiability of θ, $F!t : t : \varphi$ is realized at $\mathcal{N}(\sigma)$ in $\mathcal{K}$. So, $\mathcal{N}(\sigma) \nVdash !t : t : \varphi$. By the truth condition of $!t : t : \varphi$, (1) $t : \varphi \notin \mathcal{E}(\mathcal{N}(\sigma), !t)$ or (2) there exists v, s. t. $\mathcal{N}(\sigma)R^e v$ and $v \nVdash t\!:\!\varphi$. Here we want to show $\mathcal{N}(\sigma) \nVdash t\!:\!\varphi$. To show this, it suffices to show $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma, t)$ or there exists v, s.t. $\mathcal{N}(\sigma)R^e v$ and $v \nVdash \varphi$. From (1), by the closure condition on $\mathcal{E}$, $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$. This suffices to derive the desired disjunction $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$ or there exists v, s.t. $\mathcal{N}(\sigma)R^e v$ and $v \nVdash \varphi$. So, $\mathcal{N}(\sigma) \nVdash t\!:\!\varphi$. From (2), pick some state v_1, s.t. $\mathcal{N}(\sigma)R^e v_1$ and $v_1 \nVdash t\!:\!\varphi$. From the second part, we can obtain $\varphi \notin \mathcal{E}(v_1, t)$ or there exists v, s.t. $v_1 R^e v$ and $v \nVdash \varphi$. For the latter, we pick some state v_2, $v_1 R^e v_2$ and $v_2 \nVdash \varphi$. Since $\mathcal{N}(\sigma)R^e v_1$, by monotonicity, $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$. So, we have $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$, or $\mathcal{N}(\sigma)R^e v_1$ and $v_1 R^e v_2$ and $v_2 \nVdash \varphi$. By transitivity, the latter implies there exists v, $\mathcal{N}(\sigma)R^e v$ and $v \nVdash \varphi$. Hence, the whole statement implies $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$ or there exists v, $\mathcal{N}(\sigma)R^e v$ and $v \nVdash \varphi$. Thus, $\mathcal{N}(\sigma) \nVdash t\!:\!\varphi$. So, $Ft\!:\!\varphi$ is realized at $\mathcal{N}(\sigma)$. Therefore, $\theta \cup \{\sigma Ft\!:\!\varphi\}$ is satisfiable.

$\cdot$-rule: Suppose $\sigma Ft \cdot s : \varphi$ occurs on θ, $\cdot$-rule for $t \cdot s$ is applied and (1) $\sigma Ft\!:\!\psi \rightarrow \varphi$ is added on θ or (2) $\sigma Fs\!:\!\psi$ is added on θ (for any formula ψ).

By the assumption of satisfiability of θ, $Ft\cdot s:\varphi$ is realized at $\mathcal{N}(\sigma)$ in some $\mathcal{K}$ under some $\mathcal{N}(\sigma)$. So, $\mathcal{N}(\sigma) \nVdash t\cdot s:\varphi$. By the truth condition, we have $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t\cdot s)$ or there exists v, s.t. $\mathcal{N}(\sigma)R^e v$ and $v \nVdash \varphi$. Here we want to show $\mathcal{N}(\sigma) \nVdash t:\psi \to \varphi$ or $\mathcal{N}(\sigma) \nVdash s:\psi$. By the closure condition of $\mathcal{E}$, $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t\cdot s)$ implies either $\psi \to \varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$ or $\psi \notin \mathcal{E}(\mathcal{N}(\sigma), s)$. So, the disjunction implies (1) $\psi \to \varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$ or (2) $\psi \notin \mathcal{E}(\mathcal{N}(\sigma), s)$ or (3) there exists v, s.t. $\mathcal{N}(\sigma)R^e v$ and $v \nVdash \varphi$.

As in the previous case, for the first two cases, we can get the respective disjunct of the goal statement; however, we need a further argument to derive the goal statement from the last part. First, note that for any formula ψ, for any $\mathcal{K}$ and for any $u \in K$, $u \Vdash \psi \vee \neg\psi$. We pick some state v_1 satisfying $\mathcal{N}(\sigma)R^e v_1$ and $v_1 \nVdash \varphi$. Since $v_1 \Vdash \psi \vee \neg\psi$, $\mathcal{N}(\sigma)R^e v_1$ and $v_1 \nVdash \varphi$ and $v_1 \Vdash \psi$ or $\mathcal{N}(\sigma)R^e v_1$ and $v_1 \nVdash \varphi$ and $v_1 \nVdash \psi$. The former implies (4) there exists v, s.t. $\mathcal{N}(\sigma)R^e v$ and $v \nVdash \psi \to \varphi$ and the latter implies (5) there exists v, s.t. $\mathcal{N}(\sigma)R^e v$ and $v \nVdash \psi$.

So, we have (1) or (2) or (4) or (5). Each of (1) and (4) implies that $\psi \to \varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$ or there exists v, s.t. $\mathcal{N}(\sigma)R^e v$ and $v \nVdash \psi \to \varphi$. Each of (2) and (5) implies that $\psi \notin \mathcal{E}(\mathcal{N}(\sigma), s)$ or there exists v, s.t. $\mathcal{N}(\sigma)R^e v$ and $v \nVdash \psi$. Hence, $\mathcal{N}(\sigma) \nVdash t:\psi \to \varphi$ or $\mathcal{N}(\sigma) \nVdash s:\psi$. So, $Ft:\psi \to \varphi$ is realized at $\mathcal{N}(\sigma)$ or $Fs:\psi$ is realized at $\mathcal{N}(\sigma)$. Therefore, after applying π-rule for $t{\cdot}s$, our branch $\theta \cup \{\sigma Ft:\psi \to \varphi\}$ is satisfiable or $\theta \cup \{\sigma Fs:\psi\}$ is satisfiable.

ν-rules for $\Box$ (K, T, 4): These are the same as ordinary modal logics.

π-rule for $\Box$: The proof is similar to [57]. $\boxtimes$ (lemma)

Theorem 4.1.3 (Soundness) *If φ has a prefixed $\mathsf{S4LPN}$-tableau proof, then φ is valid in all models.*

Proof Suppose φ has an $\mathsf{S4LPN}$-tableau proof, but is not $\mathsf{S4LPN}$-valid. Say, φ does not hold at world s of some $\mathsf{S4LPN}$-model. Now a prefixed tableau begins with $1F\varphi$. Define an $\mathsf{S4LPN}$-interpretation $\mathcal{N}$ by setting $\mathcal{N}(1) = s$. Since φ is not forced at s, i.e. $s \nVdash \varphi$, the starting $\mathsf{S4LPN}$-tableau is $\mathsf{S4LPN}$-satisfiable ($\mathcal{N}(1) \nVdash (\varphi)$), so $\{1F\varphi\}$ is $\mathsf{S4LPN}$-satisfiable. By the lemma, so is every subsequent tableau. But an $\mathsf{S4LPN}$-satisfiable tableau cannot be closed, which contradicts the assumption that φ has a tableau proof. Therefore, if φ is $\mathsf{S4LPN}$-provable, then φ must be $\mathsf{S4LPN}$-valid. $\boxtimes$

We move on to the completeness theorem. Our proof is done by Lindenbaum-Henkin construction in [59]. We start from some definitions. A set S of prefixed formulas is $\mathsf{S4LPN}$-consistent if no $\mathsf{S4LPN}$-tableau for a finite part of S is closed. S is maximally $\mathsf{S4LPN}$-consistent if S is $\mathsf{S4LPN}$-consistent and no proper extension of S is $\mathsf{S4LPN}$-consistent.[7] S is π-complete provided, if $\sigma\pi \in S$, then for some integer k, $\sigma.k\pi_0 \in S$ ($\sigma.kF\varphi \in S$). S omits infinitely many integers if the set of integers that do not appear in prefixes in S is infinite.

---

[7] In this section, we use "consistent" to mean "$\mathsf{S4LPN}$-consistent," unless we explicitly note otherwise.

Lindenbaum-Henkin construction:

Enumerate all formulas in the language of $\mathsf{S4LPN}$: $\sigma_0\Phi_0, \ldots, \sigma_n\Phi_n, \ldots$.

$$\begin{cases} S_0 & = \mathrm{S}; \\ S_{n+1} & = S_n \cup \{\sigma_n\Phi_n\} \text{ if this is consistent and } \Phi_n \text{ is not } \pi; \\ & = S_n \cup \{\sigma_n\pi, \sigma_n.k\pi_0\} \text{ if } S_n \cup \{\sigma_n\Phi_n\} \text{ is consistent,} \\ & \qquad\qquad \Phi_n \text{ is } \pi \text{ and } \sigma_n.k \text{ is new;} \\ & = S_n \text{ otherwise.} \end{cases}$$

Here 'new' means that $\sigma_n.k$ does not occur in S_n or in π.

Now we state and prove a few claims.

Claim A: If S omits infinitely many integers, then S_n omits infinitely many integers.

Proof. For each S_n, there are obviously finitely many sets from S_0. At each step we put at most one new number is used. Thus, we need only finitely many integers (at most n) to construct $S_0, \ldots, S_n$. By assumption, S omits infinitely many integers. Hence, the union of the integers in S and the set of the integers used to construct $S_1, \ldots, S_n$ (a finite set of integers) must omit infinitely many integers. ⊠ (claim).

Claim B: If $S_n \cup \{\sigma_n\pi\}$ is consistent, so is $S_n \cup \{\sigma_n\pi, \sigma_n.k\pi_0\}$.

Proof. Suppose $S_n \cup \{\sigma_n\pi, \sigma_n.k\pi_0\}$ is inconsistent. Then, there is a set S^*, s.t. $S^* \subseteq S_n \cup \{\sigma_n\pi, \sigma_n.k\pi_0\}$, has a closed tableau. Note that it must be the case that $\sigma_n.k\pi_0 \in S^*$. (Otherwise, $S_n \cup \{\sigma_n\pi\}$ would be inconsistent.) Take a closed tableau of S^*. By putting the formula $\sigma_n\pi$ and apply π-rule, we can take this tableau as a tableau for $S^* \setminus \{\sigma_n.k\pi_0\}$. (Note that since k is new, $\sigma_n.k$ can be taken as a prefix produced by an application of π-rule added later.) However, then, $S^* \setminus \{\sigma_n.k\pi_0\} \subseteq S_n \cup \{\sigma_n\pi\}$. This implies $S_n \cup \{\sigma_n\pi\}$ is inconsistent. Contradiction.

Claim C: If S_n omits infinitely many integers, there will always be a new prefix.

Proof. Suppose S_n omits infinitely many integers and that $S_n \cup \{\sigma_n\pi\}$ is consistent. Choose the least integer k from the omitted integers. Since this k is new, we can construct $\sigma_n.k\pi_0$ as desired. ⊠ (claim)

Lemma 4.1.4 *If S is consistent and omits infinitely many integers, then $\bigcup_n S_n$ $(= S_\omega)$ will be maximally consistent and π-complete.*

Proof Suppose S is consistent and omits infinitely many integers. We construct S_ω following the construction above. We prove the lemma by proving the following two claims.

Claim 4.1.5 *S_ω is maximally consistent.*

Proof At first, we show S_ω is consistent. Suppose otherwise. Then, there are finite set of prefixed signed formulas $\{\sigma_1\Phi_1, \ldots, \sigma_m\Phi_m\} \subseteq S_\omega$ that has a closed tableau. But this is against the construction of S_ω. For then there must exist a formula $\sigma_n\Phi_n$ $(n \leq m)$ in this set such that $\{\sigma_1\Phi_1, \ldots, \sigma_{n-1}\Phi_{n-1}\}$ is consistent, but $\{\sigma_1\Phi_1, \ldots, \sigma_{n-1}\Phi_{n-1}, \sigma_n\Phi_n\}$ is inconsistent. But such a formula $\sigma_n\Phi_n$ must have been excluded by construction and cannot be a member of S_ω.

Next, we show that S_ω does not have any consistent proper superset of formulas. We want to show that for any $\sigma\Phi$, $\sigma\Phi \notin S_\omega$ implies $S_\omega \cup \{\sigma\Phi\}$ is inconsistent. Suppose there exists $\sigma\Phi$, s.t. $\sigma\Phi \notin S_\omega$ and $S_\omega \cup \{\sigma\Phi\}$ is consistent. The latter implies that no tableau for any finite subset $S \subseteq S_\omega \cup \{\sigma\Phi\}$ is closed. On the other hand, by the Lindenbaum-Henkin construction, $\sigma\Phi$ must occur in the fixed enumeration of formulas. So, there exists m such that $\sigma\Phi = \sigma_m\Phi_m$, where $\sigma_m\Phi_m$ is the m-th formula in the enumeration. Consider $S_m \subseteq S_\omega$ that is the set of formulas occurring at the $m+1$-st stage of the Lindenbaum-Henkin construction. Since $S_m \cup \{\sigma_m\Phi_m\} \subseteq S_\omega \cup \{\sigma_m\Phi_m\}$ and $S_m \cup \{\sigma_m\Phi_m\}$ is a finite set, any subset of $S_m \cup \{\sigma_m\Phi_m\}$ is obviously a finite subset of $S_\omega \cup \{\sigma_m\Phi_m\}$. However, since $S_\omega \cup \{\sigma_m\Phi_m\}$ is consistent by assumption, there is no closed tableau for $S_m \cup \{\sigma_m\Phi_m\}$, i.e., $S_m \cup \{\sigma_m\Phi_m\}$ is consistent. But then, by construction, $\sigma_m\Phi_m \in S_{m+1}$. Since $S_{m+1} \subseteq S_\omega$, this implies that $\sigma_m\Phi_m \in S_\omega$, so $\sigma\Phi \in S_\omega$. But this contradicts the first assumption. ⊠(claim)

Claim 4.1.6 *S_ω is π-complete.*

Proof Suppose $\sigma F\Box\varphi \in S_\omega$. Then, there exists n such that $\sigma = \sigma_n$ and $F\Box\varphi = \Phi_n$ and $S_n \cup \{\sigma_n\Phi_n\}$ is consistent. (Otherwise, $\sigma F\Box\varphi$ would not be in S_ω.) By construction, $\sigma_n.kF\varphi$ has to be in S_{n+1}. Indeed, we can show the following. By assumption, S omits infinitely many integers. So, S_{n+1} omits infinitely many integers (by claim (A)). Hence, by claim (C), it is always possible to find a new k, s.t. $\sigma_n.kF\varphi \in S_{n+1} \subseteq S_\omega$, as desired. Therefore, S_ω is π-complete. ⊠(claim)

These suffice to show the lemma. ⊠(lemma)

We construct a canonical Kripke model $\mathcal{K} = (K, R, R^e, \mathcal{E}, \mathcal{V})$ for S4LPN based on this maximal consistent set. Let K be the set of prefixes that occur in S_ω. Possible worlds will be taken to be syntactic objects, i.e., prefixes, just as in the usual Henkin construction objects in the domain are syntactic objects, i.e., terms in the language. The accessibility relations R, R^e, propositional valuation $\mathcal{V}$ and evidence function $\mathcal{E}$ are given as follows:

1. $\sigma R \sigma'$ iff σ is an initial segment of σ' ($\sigma \leq \sigma'$);
2. $\sigma R^e \sigma'$ iff $1 \leq \sigma$ and $1 \leq \sigma'$;
3. $\sigma \in \mathcal{V}(p)$ iff $\sigma T p \in S_\omega$;
4. $\varphi \in \mathcal{E}(\sigma, t)$ iff $\sigma F t{:}\varphi \notin S_\omega$.

Now we check that $\mathcal{E}$, R and R^e defined in the canonical model satisfy the conditions of a model of S4LPN. However, by construction of our canonical model, it is obvious that R is a reflexive and transitive relation and R^e is an equivalence relation and $R \subseteq R^e$. So, we focus on the conditions of $\mathcal{E}$. We first prove a useful proposition.

Proposition 4.1.7 *For any σ, σ' occurring in ${S_\omega}^8$,*

1. *$\sigma Ft{:}\varphi \in S_\omega$ if and only if $\sigma' Ft{:}\varphi \in S_\omega$*
2. *$\sigma Tt{:}\varphi \in S_\omega$ if and only if $\sigma' Tt{:}\varphi \in S_\omega$*

Proof 1. Suppose $\sigma Ft : \psi \in S_\omega$ but $\sigma' Ft : \psi \notin S_\omega$ for some $\sigma \in K$ and for some $\sigma' \in K$. By maximal consistency, $S_\omega \cup \{\sigma' Ft : \psi\}$ is inconsistent, namely there is a finite subset S_1 of $S_\omega \cup \{\sigma' Ft : \psi\}$ such that S_1 has a closed tableau. The existence of a formula Φ such that $\sigma'\Phi \in S_\omega$ is guaranteed by the condition $\sigma' \in K$. Since $\sigma' Ft : \psi \notin S_\omega$, there must be at least one prefixed signed formula that has the prefix σ' and σ' is in S_ω that is different from $\sigma' Ft : \psi$. However, the existence of a finite subset S_1 that has a closed tableau due to maximal consistency of S_ω and $\sigma' Ft : \psi \notin S_\omega$ does not guarantee that there exists a signed formula Φ, s.t. $\sigma'\Phi \in S_1$ (although $\sigma'\Phi \in S_\omega$). Nonetheless, we can make S_1 contain such a formula $\sigma'\Phi$, w.l.o.g., for the following reason. Take an arbitrary inconsistent finite subset S_1 of $S_\omega \cup \{\sigma' Ft : \psi\}$. This may or may not contain $\sigma'\Phi$. If it does, it is our desired S_1. If it does not, we can just add the formula $\sigma'\Phi$ to S_1. And use it as S_1 (only here we call it S_1'). Since $\sigma'\Phi \in S_\omega$ and S_1 that is picked before $\sigma'\Phi$ being added is already inconsistent, addition of $\sigma'\Phi$ to S_1 does not affect the issue of consistency or inconsistency of the pertinent set S_1. Hence, S_1 has a closed tableau if and only if S_1' has a closed tableau. In the following argument, w.l.o.g., we use only S_1 as a finite subset S_1 of $S_\omega \cup \{\sigma' Ft : \psi\}$ such that S_1 has a closed tableau and there exists a signed $\sigma'\Phi \in S_1$, s.t.

---

[8] Note that for any σ, $\sigma' \in S_\omega$, $1 \leq \sigma$ and $1 \leq \sigma'$ by construction.

$\sigma'\Phi \in S_\omega$ and $\Phi \neq Ft : \psi$.

However, if so, we can produce another closed tableau from the closed tableau for S_1 by taking the following steps. First, get $1Ft : \psi$ from $\sigma Ft : \psi$ by applying EF_r finitely many times. Note that σ' occurs in the tableau since it occurs in the original set S_1. Also, note that our current definition of a prefix being used plays a crucial role here. Since σ' occurs in our initial set of our tableau, we take all its initial segment of σ' occurs on the branch. Therefore, it is possible to apply EF here $1Ft : \psi$. We do that finitely many times until we get $\sigma' Ft : \psi$.

Once we obtain $\sigma' Ft : \psi$ this way, we glue a closed tableau for S_1 with the tableau just constructed. This would constitute another closed tableau whose initial set $(S_1 \backslash \{\sigma' Ft : \psi\}) \cup \{\sigma Ft : \psi\}$. Note we can explicitly construct the formula $\sigma' Ft : \psi$ by the rule EF and out of σ' existing in S_1 (since $\sigma'\Phi \in S_1$) and the formula $\sigma Ft : \psi \in S_\omega$, independently of the formula $\sigma' Ft : \psi$ which is assumed to be outside of S_ω, so that we constructed the entire closed tableau for $(S_1 \backslash \{\sigma' Ft : \psi\}) \cup \{\sigma Ft : \psi\}$. But $(S_1 \backslash \{\sigma' Ft : \psi\}) \cup \{\sigma Ft : \psi\}$ is included in S_ω. Hence, this contradicts the maximal consistency of S_ω. A proof of 2 is similar. ⊠ (proposition)

Proposition 4.1.8 *The evidence function defined above satisfies the following conditions: (1) monotonicity, (2) closure conditions, (3) constant specification.*

Proof (1) (Monotonicity) Suppose $\sigma R^e \sigma'$ and $\varphi \in \mathcal{E}(\sigma, t)$. By definition,

$\sigma F t{:}\varphi \notin S_\omega$. So, proposition 1, $\sigma' F t{:}\varphi \notin S_\omega$. So, $\varphi \in \mathcal{E}(\sigma', t)$.

(2) (Closure Conditions 2.) Suppose $\varphi \notin \mathcal{E}(\sigma, t{\cdot}s)$. By definition, $\sigma F t{\cdot}s : \varphi \in S_\omega$. We consider applying $t{\cdot}s$-rule to a finite subset of S_ω.

Claim 4.1.9 *$\sigma F t{:}\psi \to \varphi \in S_\omega$ or $\sigma F s{:}\psi \in S_\omega$.*

Proof First, we show that $S_\omega \cup \{\sigma F t{:}\psi \to \varphi\}$ is consistent or $S_\omega \cup \{\sigma F s{:}\varphi\}$ is consistent. Suppose otherwise, i.e., $S_\omega \cup \{\sigma F t : \psi \to \varphi\}$ is inconsistent and $S_\omega \cup \{\sigma F s : \varphi\}$ is inconsistent. Then, there exists a finite set S^1 s.t. $S^1 \subseteq S_\omega \cup \{\sigma F t{:}\psi \to \varphi\}$ and $\sigma F t{:}\psi \to \varphi \in S^1$ and S^1 has a closed tableau and there exists a finite set S^2 s.t. $S^2 \subseteq S_\omega \cup \{\sigma F s : \psi\}$, $\sigma F s : \psi \in S^2$ and S^2 has a closed tableau. (Note that since S_ω is consistent, we have to use an additional formula to close a tableau for each case.) These formulas are obtained by applying $\cdot$-rule. So, we can construct another closed tableau for a finite set $S^1 \cup S^2 \cup \{\sigma F t{\cdot}s : \varphi\}$ by taking the tableau for S^1 and S^2 as branches of the new tableau and by applying $\cdot$-rule. However, $S^1 \subseteq S_\omega$, $S^2 \subseteq S_\omega$ and $\{\sigma F t{\cdot}s{:}\varphi\} \subseteq S_\omega$. So, a finite subset of S_ω has a closed tableau, which contradicts the consistency of S_ω. By maximality, $\sigma F t{:}\psi \to \varphi \in S_\omega$ or $\sigma F s{:}\psi \in S_\omega$. ⊠(claim)

By definition, $\psi \to \varphi \notin \mathcal{E}(\sigma, t)$ or $\psi \notin \mathcal{E}(\sigma, s)$.

(Closure Condition 3.) Suppose $t : \varphi \notin \mathcal{E}(\sigma, !t)$. By definition, $\sigma F !t : t : \varphi \in S_\omega$. We want to show $\sigma F t{:}\varphi \in S_\omega$. To show that, it suffices to show the consistency of $S_\omega \cup \{\sigma F t : \varphi\}$, due to maximal consistency of S_ω. Suppose otherwise, i.e. this set is inconsistent. Then, there is a finite subset of this set that has a closed tableau. Pick up one such closed tableau. Put $\sigma F!t\!:\!t\!:\!\varphi$ on top of it and apply !-rule, then we can produce another closed tableau of a finite subset of $S_\omega \cup \{\sigma F!t\!:\!t\!:\!\varphi\}$. However, $\sigma F!t\!:\!t\!:\!\varphi \in S_\omega$. This implies S_ω is inconsistent. Contradiction. So, $S_\omega \cup \{\sigma Ft\!:\!\varphi\}$ is consistent and, by maximality of S_ω, $\sigma Ft\!:\!\varphi \in S_\omega$. Hence, by definition, $\varphi \notin \mathcal{E}(\sigma, t)$.

Proofs for (Closure Condition 4.) and (3) Constant Specification are similar. ⊠

Lemma 4.1.10 (Truth Lemma) *For every signed formula* Φ*,*

$\sigma\Phi \in S_\omega \Longrightarrow \sigma$ *realizes* Φ *in the model* $\mathcal{K}$*.*

Proof By induction on complexity of formulas. We show $\Box$ and $t\!:$cases.

Case 1. $\Phi = T\Box\varphi$. Suppose $\sigma T\Box\varphi \in S_\omega$. Assume, for reductio, there exists $\sigma' \in K$, s.t. $\sigma \leq \sigma'$ and $\sigma' T\varphi \notin S_\omega$. Pick an witness and temporarily call it σ'_1. Then $S_\omega \cup \{\sigma'_1 T\varphi\}$ is inconsistent. That is, there exists a finite set S_1, s.t. $S_1 \subseteq S_\omega \cup \{\sigma'_1 T\varphi\}$ and S_1 has a closed tableau. Since $\sigma'_1 \in K$, we can use the same trick to make sure that there exists a signed formula Φ, s.t. $\sigma'_1\Phi \in S_1$ and $\Phi \neq T\varphi$. Then, since $\sigma \leq \sigma'_1$, we can put $\sigma T\Box\varphi$ on top of such a tableau and construct another closed tableau by applying ν_4 finitely many times and by applying ν_T to obtain $\sigma'_1 T\varphi$ after hitting $\sigma'_1 T\Box\varphi$. (Again, the initial segments of σ'_1 are all used in this tableau by definition since we made sure that σ'_1 occurs in S_1.) Note that the whole tableau can be taken as a closed tableau for $(S_1 \setminus \{\sigma' T\varphi\}) \cup \{\sigma T\Box\varphi\}$. However, $(S_1 \setminus \{\sigma' T\varphi\}) \cup \{\sigma T\Box\varphi\} \subseteq S_\omega$. This is contradictory to maximal consistency of S_ω. Hence, for any $\sigma' \in K$, if $\sigma \leq \sigma'$ then $\sigma' T\varphi \in S_\omega$. By definition $\sigma \leq \sigma'$ iff $\sigma R \sigma'$, and by IH, $\sigma' T\varphi \in S_\omega$ implies σ' realizes φ in $\mathcal{K}$. Hence, $\sigma' \Vdash \varphi$. Thus, for any $\sigma' \in K$, if $\sigma R \sigma'$ then $\sigma' \Vdash \varphi$. Hence, $\sigma \Vdash \Box\varphi$. Therefore, σ realizes $\Box\varphi$ in $\mathcal{K}$.

Case 2. $\Phi = F\Box\varphi$. Suppose $\sigma F\Box\varphi \in S_\omega$. By π-completeness of S_ω, $\sigma.kF\varphi \in S_\omega$ for some $\sigma.k$ occurring in S_ω. By IH, $\sigma.k$ realizes $F\varphi$ in the model $\mathcal{K}$. So, $\mathcal{K}, \sigma.k \nVdash \varphi$. On the other hand, clearly, $\sigma \leq \sigma.k$, i.e., $\sigma R \sigma.k$. So, there exists σ', s.t. $\sigma R \sigma'$, $\mathcal{K}, \sigma' \nVdash \varphi$. So $\mathcal{K}, \sigma \nVdash \Box\varphi$. Hence, σ realizes $F\Box\varphi$ in the model $\mathcal{K}$.

Case 3. $\Phi = Tt : \varphi$. Suppose $\sigma Tt : \varphi \in S_\omega$. It suffices to show that $\sigma \Vdash t : \varphi$. To show this, it is suffices to show the two statements: (1) $\varphi \in \mathcal{E}(\sigma, t)$ and (2) for all σ', $(\sigma R^e \sigma' \Rightarrow \sigma' \Vdash \varphi)$. (1) is an immediate consequence of the definition of $\mathcal{E}$ and $\sigma Ft : \varphi \notin S_\omega$ (by consistency). So, $\varphi \in \mathcal{E}(\sigma, t)$. To show (2), suppose $\sigma Tt : \varphi$ and there exists $\sigma' \in K$ s.t. $\sigma R^e \sigma'$ and $\sigma' T\varphi \notin S_\omega$. Take an instance of σ' and call it σ'_1. Then $1 \leq \sigma$ and $1 \leq \sigma'_1$ and $S_\omega \cup \{\sigma'_1 T\varphi\}$ is inconsistent. Then, there exists a finite subset $S_1 \subseteq S_\omega \cup \{\sigma'_1 T\varphi\}$ s.t. S_1 has a closed tableau. By the same argument as given in the proposition 4.1.7., w.l.o.g., we can assume that $\sigma'_1\Phi$ (for some Φ s.t.$\Phi \neq T\varphi$) occurs in the set S_1. Then we can manipulate the given closed tableau: putting $\sigma Tt : \varphi$ and applying E4r (several times), we derive $1Tt : \varphi$; applying E4 (several times), we can derive $\sigma'_1 Tt : \varphi$; finally, applying ET, we can derive $\sigma'_1 T\varphi$. Then the resulting closed tableau can be taken as a closed tableau for $(S_1 \setminus \{\sigma'_1 T\varphi\}) \cup \{\sigma Tt : \varphi\}$, but this is a (finite) subset of S_ω, which is contradictory to maximal consistency of S_ω. Hence, $\sigma Tt : \varphi$ implies that for any $\sigma' \in K$, $\sigma R^e \sigma' \Rightarrow \sigma' T\varphi \in S_\omega$. By IH, σ' realizes φ in $\mathcal{K}$. So, $\sigma' \Vdash \varphi$. So, for all $\sigma' \in K$, s.t. $\sigma R^e \sigma'$, $\sigma' \Vdash \varphi$. Therefore, $\sigma \Vdash t{:}\varphi$. So, σ realizes $t{:}\varphi$ in $\mathcal{K}$.

Case 4. $\Phi = Ft{:}\varphi$. Suppose $\sigma Ft{:}\varphi \in S_\omega$. It suffices to show $\sigma \nVdash t{:}\varphi$. To show this, it is sufficient to show (1) $\varphi \notin \mathcal{E}(\sigma, t)$ or (2) there exists σ', s.t. $\sigma R^e \sigma'$ and $\sigma' \nVdash \varphi$. By the assumption, it is not the case that $\sigma Ft{:}\varphi \notin S_\omega$. So, this immediately implies $\varphi \notin \mathcal{E}(\sigma, t)$. So, $\varphi \notin \mathcal{E}(\sigma, t)$ or there exists σ', s.t. $\sigma R^e \sigma'$ and $\sigma' \nVdash \varphi$. So, $\sigma \Vdash t{:}\varphi$. Hence, σ realizes $Ft{:}\varphi$ in $\mathcal{K}$. ⊠ (Truth Lemma)

Theorem 4.1.11 (Completeness) *If φ is S4LPN-valid, then φ has a proof using the tableau rules for S4LPN.*

Proof We show the contrapositive. Suppose φ is not provable using the prefixed S4LPN-tableau rules. Then $\{1F\varphi\}$ is S4LPN-consistent, and it omits infinitely many integers. So, we can extend it to a maximally S4LPN-consistent, π-complete set S_ω by the above construction and the lemma 4.1.4. We can define a canonical Kripke model $\mathcal{K}$ out of S_ω. By Truth Lemma, we can show $F\varphi$ is realized at 1 in $\mathcal{K}$. So, there is a Kripke model $\mathcal{K}$ and a state σ such that $\mathcal{K}, \sigma \nVdash \varphi$. ⊠

Note that this proof of completeness does not use cut anywhere. So, the prefixed tableau system for S4LPN is a complete cut-free system.

4.1.3 Hypersequent Calculus for S4LPN

Here we first present the hypersequent calculus HS4LPN. And then we give a translation from the prefixed tableau system to the hypersequent calculus.

In our formulation in this paper, a hypersequent is a finite set of sequents in traditional Gentzen-style sequent calculi, which is written as follows. $\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_n \Rightarrow \Delta_n$, where "|" has a disjunctive reading. (The precise definition of this will be given later.) To make our translation from the prefixed tableau system smooth, we take a sequent to be a pair of finite sets of formulas. So, a hypersequent is a set of (pairs of) sets of formulas. Because of this formulation, some structural rules, i.e., internal and external exchange rules, internal and external contraction rules, can be omitted from our systems.

1) **Axiom**: $A \Rightarrow A$ $\quad \bot \Rightarrow$

2) **External structural rules**: $\textbf{EW}\ \dfrac{G}{G|H}$

3) **Internal structural rules**:

$$\textbf{LW}\ \frac{G|\Gamma \Rightarrow \Delta}{G|A, \Gamma \Rightarrow \Delta} \qquad \textbf{RW}\ \frac{G|\Gamma \Rightarrow \Delta}{G|\Gamma \Rightarrow \Delta, A}$$

4) **Operational rules**:

$$\mathbf{L}\wedge\ \frac{G|A,B,\Gamma\Rightarrow\Delta}{G|A\wedge B,\Gamma\Rightarrow\Delta}\qquad\qquad \mathbf{R}\wedge\ \frac{G|\Gamma\Rightarrow\Delta,A\qquad G|\Gamma\Rightarrow\Delta,B}{G|\Gamma\Rightarrow\Delta,A\wedge B}$$

$$\mathbf{L}\vee\ \frac{G|A,\Gamma\Rightarrow\Delta\quad G|B,\Gamma\Rightarrow\Delta}{G|A\vee B,\Gamma\Rightarrow\Delta}\qquad\qquad \mathbf{R}\vee\ \frac{G|\Gamma\Rightarrow\Delta,A,B}{G|\Gamma\Rightarrow\Delta,A\vee B}$$

$$\mathbf{L}\rightarrow\ \frac{G|\Gamma\Rightarrow\Delta,A\qquad G|B,\Gamma\Rightarrow\Delta}{G|A\rightarrow B,\Gamma\Rightarrow\Delta}\qquad \mathbf{R}\rightarrow\ \frac{G|A,\Gamma\Rightarrow\Delta,B}{G|\Gamma\Rightarrow\Delta,A\rightarrow B}$$

$$\mathbf{L}\neg\ \frac{G|\Gamma\Rightarrow\Delta,A}{G|\Gamma,\neg A\Rightarrow\Delta}\qquad\qquad \mathbf{R}\neg\ \frac{G|\Gamma,A\Rightarrow\Delta}{G|\Gamma\Rightarrow\Delta,\neg A}$$

5) **Rules for S4□:** $\mathbf{L}\Box\ \frac{G|A,\Gamma\Rightarrow\Delta}{G|\Box A,\Gamma\Rightarrow\Delta}\qquad \mathbf{R}\Box\ \frac{G|\Box\Gamma\Rightarrow A}{G|\Box\Gamma\Rightarrow\Box A}$

6) **Rules for Proof-terms of LP:**

$$\mathbf{Lt}\ \frac{G|A,\Gamma\Rightarrow\Delta}{G|t{:}A,\Gamma\Rightarrow\Delta}\qquad \mathbf{R}\cdot\ \frac{G|\Gamma\Rightarrow\Delta,t{:}A\rightarrow B\qquad G|\Gamma\Rightarrow\Delta,s{:}A}{G|\Gamma\Rightarrow\Delta,t{\cdot}s{:}B}$$

$$\mathbf{R}!\ \frac{G|\Gamma\Rightarrow\Delta,t{:}A}{G|\Gamma\Rightarrow\Delta,!t{:}t{:}A}\qquad \mathbf{R}+\ \frac{G|\Gamma\Rightarrow\Delta,t{:}A}{G|\Gamma\Rightarrow\Delta,t+s{:}A}\qquad \frac{G|\Gamma\Rightarrow\Delta,s{:}A}{G|\Gamma\Rightarrow\Delta,t+s{:}A}$$

7) **Constant Specification Rule**:

$\dfrac{}{\Rightarrow c{:}A}$ where A is an axiom of $\mathsf{S4LPN}$ and $A \in \mathcal{CS}(c)$.

8) **Labeled Splitting**[9]:
$$\frac{G|\vec{t}{:}\Gamma_1, \Gamma_2 \Rightarrow \vec{s}{:}\Delta_1, \Delta_2}{G|\vec{t}{:}\Gamma_1 \Rightarrow \vec{s}{:}\Delta_1|\Gamma_2 \Rightarrow \Delta_2}$$

9) **Cut**:
$$\frac{G_1|\Gamma_1 \Rightarrow \Delta_1, A \qquad G_2|A, \Gamma_2 \Rightarrow \Delta_2}{G_1|G_2|\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2}$$

E.g., a derivation in $\mathsf{HS4LPN}$.

$$\begin{array}{c}
x{:}A \Rightarrow x{:}A \qquad c{:}(A \to (B \to A)) \\
\hline
x{:}A \Rightarrow c\cdot x{:}B \to A \\
\hline
x{:}A \Rightarrow | \Rightarrow !(c\cdot x){:}c\cdot x{:}B \to A \\
\hline\hline
\Rightarrow \neg x{:}A|\neg(!(c\cdot x){:}c\cdot x{:}B \to A) \Rightarrow \\
\hline
\Rightarrow \Box\neg x{:}A|\neg(!(c\cdot x){:}c\cdot x{:}B \to A) \Rightarrow \\
\hline\hline
\neg(!(c\cdot x){:}c\cdot x{:}B \to A) \Rightarrow \Box\neg x{:}A|\neg(!(c\cdot x){:}c\cdot x{:}B \to A) \Rightarrow \Box\neg x{:}A \\
\hline
\neg(!(c\cdot x){:}c\cdot x{:}B \to A) \Rightarrow \Box\neg x{:}A
\end{array}$$

(The first double line corresponds to a few operational rules and the second corresponds to a few internal weakening.)

---

[9] Here $\vec{t}:\Gamma_1 = t_1 : \varphi_1, \ldots, t_n : \varphi_n$. The rule covers cases where $\vec{t}:\Gamma_1$ or $\vec{s}:\Delta_1$ is empty ([18]). Note also that this rule has its origin in Avron's modal splitting rule in his hypersequent calculus for S5, which can handle the negative introspection in S5.

By the following translation from hypersequents to formulas in the language of S4LPN, we can prove that the Hilbert-style system S4LPN and HS4LPN are deductively equivalent.

$$\mathcal{I}(\Gamma_1 \Rightarrow \Delta_1 | \dots | \Gamma_n \Rightarrow \Delta_n) = \Box(\bigwedge \Gamma_1 \to \bigvee \Delta_1) \vee \dots \vee \Box(\bigwedge \Gamma_n \to \bigvee \Delta_n)$$

Theorem 4.1.12 *HS4LPN*$\vdash \Gamma_1 \Rightarrow \Delta_1 | \dots | \Gamma_n \Rightarrow \Delta_n$ *if and only if S4LPN*$\vdash$ $\Box(\bigwedge \Gamma_1 \to \bigvee \Delta_1) \vee \dots \vee \Box(\bigwedge \Gamma_n \to \bigvee \Delta_n)$.

Proof In both ways, a proof is given by induction on the length of derivation. To derive Labeled Splitting (from right to left), we can use the fact S4LPN$\vdash$ $(t_1 : A \to t_2 : B) \to \Box(t_1 : A \to t_2 : B)$. ⊠

4.1.4 Translation from the prefixed tableau system to the hypersequent calculus

The theorem stating the deductive equivalence between the Hilbert-style axiom system and the hypersequent calculus does not tell us whether cut is admissible in our hypersequent calculus. In order to show that cut is admissible, we need another argument. Here we give a semantic proof of cut-admissibility. Since the Hilbert-style system S4LPN is sound and HS4LPN is equivalent to that system, HS4LPN is sound with respect to the Kripke semantics given above. So, to show cut-admissibility, it suffices to show completeness of HS4LPN without cut (HS4LPN$^-$) with respect to the same semantics. We show completeness of $\mathsf{HS4LPN}^-$ by providing a way of translating a prefixed tableau proof to a proof in the hypersequent calculus, following [59].

4.1.4.1 Definition of Translation

We now define a translation from the language of the prefixed tableau system $\mathsf{TS4LPN}$ to that of the hypersequent calculus $\mathsf{HS4LPN}$. By using that, we show that the derivability in the prefixed tableau system is preserved under the translation in the hypersequent calculus. Our translation mapping, which is called s, is defined in two stages.

First, we define a mapping t that maps a set of prefixed formulas to a set of sets of signed formulas in the following way.

1. The set of prefixed formulas is partitioned into subsets so that all formulas with the same prefixes σ_i go into the same subset.

2. We strip off prefixes from those partitioned prefixed formulas (for each σ_i).

3. We call the resulting set H_{σ_i} for each σ_i, i.e. $H_{\sigma_i} := \{\Phi | \sigma_i \Phi \in S\}$.

4. We arrange these H_{σ_i}'s by using some order[10] in the form of a hypersequent via "|". So, we have $H_1|\dots|H_{\sigma_1}|H_{\sigma_2}|\dots|H_{\sigma_n}$. Our reading "|" is the same as that of hypersequent, so we have now constructed a set of sets of signed formulas.[11] Let $S^t := H_1|\dots|H_{\sigma_1}|H_{\sigma_2}|\dots|H_{\sigma_n}$.

---

[10] The order can be arbitrary since our hypersequents are sets of sets of formulas.

[11] Each H_{σ_i} will work as each sequent occurring in the hypersequent that we will obtain as an image of the mapping s we are defining. But note that each H_{σ_i} consists of only

Second, we consider a mapping that maps a set of sets of signed formulas, i.e. $H_1| \ldots \ |H_{\sigma_1}|H_{\sigma_2}| \ldots |H_{\sigma_n}$, to a set of sequents. This mapping can be readily constructed by putting T formulas to the antecedent and F formulas to the succedent for each case of H_{σ_i}. Namely, if $H_{\sigma_i} = \{T\varphi_1, \ldots, T\varphi_k, F\psi_1, \ldots, F\psi_m\}$, then we map this to $\varphi_1, \ldots, \varphi_k \Rightarrow \psi_1, \ldots, \psi_m$. This is the only thing that we have to do for each H_{σ_i}, but, for notational simplicity, we officially define mapping for an entire set of H_{σ_i}'s. We call this mapping u.

$(H_1| \ldots |H_{\sigma_1}|H_{\sigma_2}| \ldots |H_{\sigma_n})^u = \varphi_{1,1}, \ldots, \varphi_{1,k_{\sigma_1}} \Rightarrow \psi_{1,1}, \ldots, \psi_{1,m_{\sigma_1}}|\varphi_{\sigma_1,1}, \ldots, \varphi_{\sigma_1,k_{\sigma_1}} \Rightarrow \psi_{\sigma_1,1}, \ldots, \psi_{\sigma_1,m_{\sigma_1}}| \ldots |\varphi_{\sigma_n,1}, \ldots, \varphi_{\sigma_n,k_{\sigma_n}} \Rightarrow \psi_{\sigma_n,1}, \ldots, \psi_{\sigma_n,m_{\sigma_n}}$.

We finally define the desired mapping s as a composition of t and u, i.e., $S^s := ((S^t)^u)$. So, S^s will be of the form $\Gamma_1 \Rightarrow \Delta_1| \ldots |\Gamma_{\sigma_i} \Rightarrow \Delta_{\sigma_i}| \ldots |\Gamma_{\sigma_n} \Rightarrow \Delta_{\sigma_n}$.

4.1.4.2 Potential problems and an outline of their solution

We need some special care in order to guarantee that the translation defined above from the prefixed tableau system for S4LPN to the hypersequent calculus for S4LPN preserves provability.

To see what care is sufficient, we start from some general observations concerning the differences of prefixed tableau systems and hypersequent calculi.

(1) Prefixed tableau systems have flexibility in the order of applications

---

signed formulas.

of rules.

(2) HS4LPN has a modal rule that is essentially the same as R$\Box$ rule in a sequent calculus for S4. Since this is a counterpart of π-rule in a destructive tableau system, the provability in HS4LPN depends on the order of applications of rules.

(3) Also, there is no rule in HS4LPN which exactly corresponds to ν-rules for K or 4 (We write $\nu_{K,4}$-rule for ν-rules for K and 4). The (hyper)sequent rule roughly corresponding to $\nu_{K,4}$ is R$\Box$ for S4, but R$\Box$ is different from $\nu_{K,4}$. It is not formulated separately from π.

Such differences can raise problematic cases, which can be classified into the one involving E4r or EFr (we call these "reverse rules") and the one that does not.

i) π-rule is applied earlier than α, β-rule(s) in a prefixed tableau, so in the translated system, we miss some cases of possible applications of $\Box R$ corresponding to ν-rule with new prefixes (since the outermost logical symbols of the relevant formulas are not $\Box$.) (E.g. $\{1T\Box A \vee \Box B, 1F\Box(A \vee B)\}$.)

ii) After π-rule is applied, some formula(s) with proof terms move(s) back to previous worlds by reverse rules, and $\nu_{K,4}$ rules are applied subsequently.

Here is a counterexample against a naive translation of TS4LPN, due to ii).[12]

$1F\Box(t : \Box A \to A)$

---

[12] To save space, we use ";" to write more than one formula horizontally.

$1.1.Ft : \Box A \to A$;

$1.1.Tt : \Box A$; $1.1.FA$

$1Tt : \Box A$; $1T\Box A$

1.1TA

×

One of step of the translation of this prefixed tableau proof is given as follows.

$$\frac{t : \Box A, \Box A \Rightarrow \Box(t : \Box A \to A) | t : \Box A, A \Rightarrow t : \Box A \to A, A}{t : \Box A, \Box A \Rightarrow \Box(t : \Box A \to A) | t : \Box A \Rightarrow t : \Box A \to A, A} ?$$

Since there is no rule exactly corresponding to ν_K, there is no way of directly justifying this step in HS4LPN. Hence, it is definitely not the case that any prefixed proof can be translated into a hypersequent proof.[13] The problem is raised due to the combination of an application of a reverse rule and subsequent applications of $\nu_{K,4}$, since if we apply only a reverse rule then it can be easily handled by Labeled Splitting in $\mathsf{HS4LPN}^-$. A more serious problem is raised by subsequent application of $\nu_{K,4}$. We have two sources of the problem here. 1) Reverse rules produce a context in which $\nu_{K,4}$ are applied separately from π. 2) Reverse rules makes it happen that a prefix is used more than once when we apply $\nu_{K,4}$.

Nonetheless, it may still be that any prefixed tableau proof can be transformed into some prefixed tableau proof that is translatable into a hyper-

---

[13] In this sense, the previous version of the proof in [98] has a gap. The current proof fills the gap.

sequent. Indeed, this is what we show in order to solve the problems. To handle particular problems raised by i) and ii) 1), it suffices to use "proof confluence property" of prefixed tableau systems, which allows reordering of applications of rules in a prefixed tableau proof. For ii) 2), we use preservation of closure under taking numerical variants of tableaux in addition to proof confluence.

However, only by using the manipulation of a given tableau proof, the discrepancy pointed out in (3) is not entirely resolved (applications of $\nu_{K,4}$). To resolve that, it suffices to modify the prefixed tableau system so that the relevant applications of $\nu_{K,4}$ become essentially redundant. Namely, we modify π-rule as follows.

$\pi^\sharp$-rule: Let S be a set of prefixed formulas (a tableau constructed up to the previous step). Let $\sigma F \Box\varphi \in S$ be the premise, let $\sigma.nF\varphi$ be the conclusion of the traditional π-rule ($\sigma.n$ is new on a branch), and let $S^\sharp_{\sigma.n} = \{\sigma.nT\Box\psi, \sigma.nT\psi | \sigma T\Box\psi \in S\}$.

$$\frac{\{\sigma F\Box\varphi\} \cup S}{\{\sigma.nF\varphi\} \cup S^\sharp_{\sigma.n}}$$

This rule combines the idea of constructing a new prefix in the prefixed tableau and the idea of destructive tableau. This rule carries $\Box\psi, \psi$ to the new world automatically. (Since a variant of the prefixed tableau system (an auxiliary system for translation) that we eventually adopt spoils "strong proof confluence”, we treat this rule not as a primitive rule in TS4LPN, but we introduce a variant of TS4LPN in which $\pi^\sharp$ is treated as primitive. We need to use proof confluence before we introduce the auxiliary system is introduced.)

Combining these ideas, we translate a proof TS4LPN into a proof in HS4LPN. Our immediate goal is to obtain a method of manipulating an arbitrary (cut-free) proof in TS4LPN into a proof satisfying the condition: all applications of $\nu_{K,4}$-rules are made immediately after an application of π-rule where the application of $\nu_{K,4}$ uses the prefix introduced by the π rule preceding it. Then we move on to $\pi^\sharp$.

4.1.4.3 Proof confluence in prefixed tableau systems

Here we prove proof confluence for TS4LPN, namely a version of the Church-Rosser property of a single-step prefixed tableau system (modulo renaming of prefixes). Massacci [102] briefly explains the notion as follows, “loosely speaking, confluence means that the order in which we select the rules does not substantially matter: we can always “converge” to the same result without backtracking.” (p.323, [102])

Now we give some definitions ([82], [41]). Let us use x, y, z to stand for stages of computation and $\longrightarrow$ to stand for a relation between them. In our context, the computation is the proof search in the prefixed tableau system, and its stages are tableaux. $x \longrightarrow y$ if the tableau y is obtained from x with an application of a tableau rule (a single step of reduction in the sense of a term rewriting systems). The relation $\longrightarrow^*$ is the reflexive and transitive closure of $\longrightarrow$, and $\longrightarrow^\epsilon$ is the reflexive closure. s and t are joinable (and we write $s \downarrow t$) if $\exists u(s \longrightarrow^* u$ and $t \longrightarrow^* u)$. s and t are meetable (and we write $s \uparrow t$) if $\exists r(r \longrightarrow^* s$ and $r \longrightarrow^* t)$.

We define an equivalence relation between stages of computation based on renaming of prefixes. In order to do that, let us first define renaming function h for prefixes ([102]) and then prove a proposition.

Definition 4.1.13 *An injective and surjective function h from the set of prefixes onto itself is a renaming if and only if $h(1) = 1$ and $h(\sigma.n) = h(\sigma).m$ for some integer m.*

Proposition 4.1.14 *Let $\mathcal{B}_i$ be the branch of its numerical variants be $\mathcal{B}_i$ with respect to h. A rule (r) can be applied to $\sigma\Phi$ in $\mathcal{B}_i$ if and only if it can be applied to $h(\sigma)\Phi$ in $\mathcal{B}_j$.*

Proof For cases in which we do not change the prefix of the premise in applying the rule, applicability of a rule trivially holds. So, we only need to check $\nu_{K,4}$ cases (explicit versions with proof-terms are similar) and π-rules. Note that the domain of the mapping h is as follows: $Dom(h) = \{\sigma | \sigma\Phi \in \mathcal{B}_i$ for some Φ $\}$ where $\mathcal{B}_i$ is a branch in the original tableau.

ν_K-rule (ν_4 case is similar) : Suppose $\sigma T\Box\varphi \in \mathcal{B}_i$ and $\mathcal{B}_i$ satisfies the precondition for applying ν_K rule, namely $\sigma.n\Phi \in \mathcal{B}_i$ for some Φ. The first conjunct implies $h(\sigma)T\Box\varphi \in \mathcal{B}_j$. Also, the second conjunct implies that $\sigma.n \in Dom(h)$. Hence, $h(\sigma.n)\Phi \in \mathcal{B}_j$.

π-rule for S4 : Suppose $\sigma F \Box \varphi \in \mathcal{B}_i$ and $\mathcal{B}_i$ satisfies the precondition for applying π-rule, namely $\sigma.m\Phi \notin \mathcal{B}_i$ for any Φ. The first conjunct immediately implies $h(\sigma)F\Box\varphi \in \mathcal{B}_j$. To derive the desired second conjunct, suppose $h(\sigma.m)\Phi \in \mathcal{B}_j$ for some Φ. Then the mapping h has the image of $\sigma.m$, i.e., $\sigma.m \in Dom(h)$. Thus, $\sigma.m\Phi \in \mathcal{B}_i$ for some Φ. But this contradicts the assumption. (The other directions are similar since h is a bijection.) ⊠

Then we can define an equivalence relation $\sim$ between stages of computation.

Definition 4.1.15 *The set of prefixed formulas $\mathcal{B}_1$ and $\mathcal{B}_2$ are equivalent modulo a renaming of prefixes if and only if there are two renamings h_{12} and h_{21} such that $h_{ij}(h_{ji}(\sigma)) = \sigma$ and if $\sigma\Phi \in \mathcal{B}_i$ then $h_{ij}(\sigma)\Phi \in \mathcal{B}_j$ for $i, j = 1, 2$.*

The definition can be extended to a tableau as sets of branches. Based on these definitions, we can state the definition of strong confluence and confluence.

Definition 4.1.16 ([102]) [14]

The relation $\longrightarrow$ is strongly confluent modulo $\sim$ if and only if $\forall x_1, x_2, y_1, y_2$, if $x_1 \sim x_2$ and $x_1 \longrightarrow y_1$ and $x_2 \longrightarrow y_2$, then $\exists u_1, u_2$, s.t. $u_1 \sim u_2$ and $y_1 \longrightarrow^* u_1$ and $y_2 \longrightarrow^\epsilon u_2$.

---

[14] In [82], another notion "locally confluent" is defined as follows. if $x \longrightarrow y_1$ and $x \longrightarrow y_2$, then $\exists u$, s.t. $y_1 \longrightarrow^* u$ and $y_2 \longrightarrow^* u$. Strong confluence is stronger than this. For the local confluence to imply confluence, a strong condition such as noetherian (there is no infinite sequence of one-step reductions) is required, but the strong confluence imply confluence due to the following lemma: if $x \longrightarrow y_1$ and $x \longrightarrow^* y_2$, then $y_1 \downarrow y_2$. That is why we use the notion of strong confluence. Hence, ϵ is used in one of the arrows in the definition. (We have omitted $\sim$ here since that is irrelevant to the difference between strong and local confluence.)

Definition 4.1.17 (Confluence [41]) *A binary relation is confluent if any two elements are joinable when they are meetable* ($\uparrow\subseteq\downarrow$)

Theorem 4.1.18 (Strong Proof Confluence) *If ν-formulas (also t : formulas in our case) can be reduced more than once, then single step tableau rules are strongly confluent modulo renaming of prefixes.*

The theorem implies proof confluence, namely that provability of a single step prefixed tableau system does not depend on the order of applications of tableau rules whenever we have a choice. Note, however, that Massacci remarks that proof confluence does not imply arbitrary shuffling of rules. There is no way of modifying the order of applications in the following cases (here $x \gtrdot y$ means "x precedes y" in the order of applications on a relevant branch): (1) $\pi \gtrdot \nu_{K,4}$; (2) $\pi \gtrdot E4r(orEFr) \gtrdot ET(orEK) \gtrdot \nu_{K,4}$, s.t. $\nu_{K,4}$ use the prefix introduced by the relevant application of π at issue here.[15] Proof confluence makes sense only if we have some choice about applications of rules.

Massacci [102] proves a stronger statement which immediately implies theorem 4.1.18.

---

[15] There are other cases where the order of applications of rules cannot be changed. 1. One rule is applied to a (proper) subformula of the formula to which the other rule is applied, 2. explicit rules also have to be applied only after some application of π. However, unlike the cases involving $\nu_{K,4}$, these cases of fixed order of applications of rules raise no problem in the translation. Case 1 is not a problem unless it is related to a change of a prefix, since a case with no change of a prefix works similarly to a sequent rule. But then the problem is essentially reduced to one where π is involved. Case 2 is not a problem since formulas with proof-terms themselves raise no problem in HS4LPN unless subsequently we apply $\nu_{K,4}$. (But the cases involving prefixes are special cases of (2) above.) Such cases have already been discussed above.

Lemma 4.1.19 (One-step confluence (modulo renaming)) *$\forall x, x', y, z$, if $x \sim x'$ and $x \longrightarrow y$ and $x' \longrightarrow z$, then $\exists u, u_e$ ($y \longrightarrow u$ and $z \longrightarrow u_e$ and $u \sim u_e$).*

Tableau rules are also formulated as reduction rules. E.g.,

Reduction precondition	Reduction relation
$\alpha : \sigma TA \wedge B \in \mathcal{B}$	$\mathcal{B} \longrightarrow_\alpha \mathcal{B} \cup \{\sigma TA, \sigma TB\}$.
$\beta : \sigma TA \vee B \in \mathcal{B}$	$\mathcal{B} \longrightarrow_\beta \mathcal{B} \cup \{\sigma TA\} \vert \mathcal{B} \cup \{\sigma TB\}$.
$\nu : \sigma T\Box A \in \mathcal{B}$ and $\exists\Phi$, s.t. $\sigma.n\Phi \in \mathcal{B}$	$\mathcal{B} \longrightarrow_K \mathcal{B} \cup \{\sigma.nTA\}$.
$\pi : \sigma F\Box A \in \mathcal{B}$ and $\forall\Phi$, s.t. $\sigma.n\Phi \notin \mathcal{B}$	$\mathcal{B} \longrightarrow_\pi \mathcal{B} \cup \{\sigma.nFA\}$.

Massacci uses Knuth-Bendix method, which proves the critical pair lemma [102]. (Here a critical pair "consists of two ways in which the common instance reduces by the two rules" ([82]). We refer to [82] and [41] for a formal definition of the notion, which involves the notion of mgu, but in our context, we do not have free variables, so we do not have to consider substitution. Hence, our case is like a reduction in ground terms. Also, the word "superposition" stands for a process of producing critical pairs.) For single step tableaux, a critical pair can only be formed when we reduce two formulas on the same branch, since reductions in different branches do not interact. Massacci begins a proof of the lemma with an observation that formulas with different prefixes do not interact each other (whose proof is omitted here, since it is straightforward).

Proposition 4.1.20 *Let* $\sigma_1\Phi_1$ *and* $\sigma_2\Phi_2$ *be prefixed signed formulas with* $\sigma_1 \neq \sigma_2$ *and let* $\mathcal{B}'$ *be the reduction of* $\mathcal{B}$ *using rule* (r_1) *on* $\sigma_1\Phi_1$*. If rule* (r_2) *can be applied to* $\sigma_2\Phi_2$ *in* $\mathcal{B}$*, then it can be applied in* $\mathcal{B}'$*.*

Proof of lemma 4.1.19 We follow Massacci's presentation of a proof in giving the details of the case of the rule for K and a case where $x = x'$ and explaining how to extend it to more general cases.

Due to the proposition, the cases of superposition are reduced to the four cases: (1) $\sigma T\Box A$, $\sigma.n\Phi_1$ and $\sigma.m\Phi_2$ are on $\mathcal{B}$; (2) $\sigma T\Box A$, $\sigma T\Box C$ and $\sigma.n\Phi$ are on $\mathcal{B}$; (3) $\sigma T\Box A$, $\sigma F\Box C$ and $\sigma.n\Phi_1$ are on $\mathcal{B}$ but no prefixed formula $\sigma.m\Phi_2$ is present on $\mathcal{B}$ ($\sigma.m$ is new); (4) $\sigma F\Box A$, $\sigma F\Box C$ are on $\mathcal{B}$ and the prefixes $\sigma.n$ and $\sigma.m$ are new. (And we also have their explicit counterparts.)

(I) **Extension of this proof to other logics**: These cases for K can be extended to other logics by replacing the prefixed formulas $\sigma T\Box A$, $\sigma.nTA$ for ν_K-rule with the premise and the conclusion of each ν-rule (and explicit rules). Each ν-formula (and t : formula) must be reducible more than once under appropriate preconditions because each logic requires more than one ν-rule and all such rules must be applicable. Due to formulas of the form $t : \varphi$, we have to check the following cases (1)-(4).

Case (1) $\nu_{K,4}$, ν_T (for $\Box$), ET, EK, E4, EF, E4r, EFr, $\cdot$-rule, !-rule, +-rule (for explicit cases) can possibly be applied in this case with the appropriate precondition ($\sigma \neq 1$) satisfied in the reverse rules. However, for ν_T, ET, E4r, EFr, $\cdot$-rule and !-rule, there is no choice. Thus, these cases are trivial.

1.1. $\sigma T \Box A$, $\sigma.n\Phi_1$, $\sigma.m\Phi_2$ are on $\mathcal{B}$ (use $\nu_{K,4}$)

1.2. $\sigma Tt : A$ (or $\sigma Ft : A$), $\sigma.n\Phi_1$, $\sigma.m\Phi_2$ are on $\mathcal{B}$ (use ET, EK, E4, EF, $+$-rule[16].)

Proof of case (1): (Only for K case. Other cases are similar.) Suppose we apply ν_K-rule either for $\sigma.n$ or $\sigma.m$. Suppose we apply ν_K-rule to $\sigma.n$ yielding $\mathcal{B} \longrightarrow \mathcal{B} \cup \{\sigma.nTA\}$. Since $\sigma.m\Phi_2$ is still present on the new branch and ν-formula can be used again, we apply ν_K-rule and obtain $\mathcal{B} \cup \{\sigma.nTA, \sigma.mTA\}$. Applying ν_K-rule first to $\sigma.m$ and then to $\sigma.n$, we obtain $\mathcal{B} \cup \{\sigma.mTA, \sigma.nTA\}$. The result is the same.

Case (2) Pick any two elements from $\{\sigma T\Box\varphi, \sigma T\Box\psi, \sigma Tt_1 : \varphi_1, \sigma Ts_1 : \psi_1, \sigma Ft_2 : \varphi_2, \sigma Fs_2 : \psi_2\}$ and use all the combinations of $\nu_{K,4}$, ν_T and EK, E4, EF with the precondition that $\sigma.n\Phi \in \mathcal{B}$, and E4r and EFr with the precondition that $\sigma \neq 1$, and operational rules (for ν_T and operational rules, we have no precondition for prefix).

Proof of case (2): Use the same argument for K as above (others are similar). The final outcome of both reduction paths is equal to $\mathcal{B} \cup \{\sigma.nTA, \sigma.nTC\}$.

Case (3) 3.1. $\sigma T\Box A$ (apply $\nu_{K,4}$, ν_T), $\sigma F\Box C$ and $\sigma.n\Phi_1$ are on $\mathcal{B}$ but no prefixed formula $\sigma.m\Phi_2$ is present on $\mathcal{B}$ ($\sigma.m$ is new).

---

[16] This introduces a choice without the precondition for prefix. So, this can be taken as a special case of 1.2.

3.2. $\sigma Tt : A$ (apply ET, EK, E4, EF, E4r) or $\sigma Ft : A$ (apply EF, EFr, or any operational rule), $\sigma F\Box C$ and $\sigma.n\Phi_1$ are on $\mathcal{B}$ (or $\sigma \neq 1$ for E4r or EFr, or no precondition for prefix in ET) but no formula $\sigma.m\Phi_2$ is present on $\mathcal{B}$ ($\sigma.m$ is new).

Proof of case (3): (Again we show only K.) We can use ν_K-rule or use π-rule and introduce a new prefix. If we use ν_K-rule, we do not introduce any new prefix and $\sigma.m$ would still be new in $\mathcal{B} \cup \{\sigma.nTA\}$. Thus, applying π-rule, we obtain $\mathcal{B} \cup \{\sigma.nTA, \sigma.mFC\}$. If we use π-rule first, then we obtain $\mathcal{B} \cup \{\sigma.mFC\}$. By assumption, for all Φ, $\sigma.m\Phi \notin \mathcal{B}$. Since $\sigma.nTD \in \mathcal{B}$, $\sigma.m \neq \sigma.n$. So, we can apply ν_K-rule and obtain $\mathcal{B} \cup \{\sigma.nTA, \sigma.mFC\}$.

Proof of case (4): (Cases are not increased here.) We must use renaming to prove confluence in this case. Suppose we reduce first $\sigma F\Box A$ and obtain $\mathcal{B} \cup \{\sigma.n_1FA\}$. In the new branch, the prefix $\sigma.n_1$ is no longer new. So, the next reduction forces the use of $\sigma.m_1$. Then we get $\mathcal{B}_1 = \mathcal{B} \cup \{\sigma.n_1FA, \sigma.m_1FC\}$. If we reduce first $\sigma F\Box C$ with $\sigma.n_2$ and obtain $\mathcal{B}\cup\{\sigma.n_2FC\}$. By another π-reduction, we get $\mathcal{B}_2 = \mathcal{B}\cup\{\sigma.n_2FA, \sigma.m_2FC\}$. At this stage, we can only guarantee that there are two new prefixes $\sigma.n$ and $\sigma.m$. It can be that $n_1 = n \neq m = n_2$ and $A \neq C$. Thus, $\mathcal{B}_1 \neq \mathcal{B}_2$. Then we define two renamings h_{12} and h_{21}.

$$h_{ij}(s) = \begin{cases} \sigma.n_j & \text{if } s = \sigma.n_i \\ \sigma.m_j & \text{if } s = \sigma.m_i \\ \text{s} & \text{otherwise} \end{cases}$$

Now we can prove that 1. $\sigma\Phi \in \mathcal{B}_i \Longrightarrow h_{ij}(\sigma)\Phi \in \mathcal{B}_j$ and 2. $h_{ij}(h_{ji}(\sigma)) = \sigma$.

Note: The use of operational rules produces no substantial difference in our proof. We present $\cdot$ case (with EF), since this rule makes the greatest difference from Massacci's single-step tableaux. Given $\sigma Ft \cdot s : A \in \mathcal{B}$ and $\sigma.n\Phi \in \mathcal{B}$ (or $\sigma \neq 1$ for reverse rules), there two ways of applying rules. The end result is clearly the same.[17]

1) $\mathcal{B} \longrightarrow_{EF} \mathcal{B} \cup \{\sigma.nFt \cdot s : A\} \longrightarrow_{t \cdot s} \mathcal{B} \cup \{\sigma.nFt \cdot s : A, \sigma Ft : B \rightarrow A\} | \mathcal{B} \cup \{\sigma.nFt \cdot s : A, \sigma Ft \cdot s : A, \sigma Fs : B\}$.
2) $\mathcal{B} \longrightarrow_{t \cdot s} \mathcal{B} \cup \{\sigma Ft : B \rightarrow A\} | \mathcal{B} \cup \{\sigma Fs : B\} \longrightarrow_{EF} \mathcal{B} \cup \{\sigma.nFt \cdot s : A, \sigma Ft : B \rightarrow A\} | \mathcal{B} \cup \{\sigma.nFt \cdot s : A, \sigma Fs : B\}$.

(II) **Extension of the proof to a general case where $x \sim x'$**

The proof of the general case $x \sim x'$ follows the same pattern, since we have two branches $\mathcal{B}_1 \sim \mathcal{B}_2$ and two renamings $h_{12} : \mathcal{B}_1 \longrightarrow \mathcal{B}_2$ and $h_{21} : \mathcal{B}_2 \longrightarrow \mathcal{B}_1$. The proposition for the general case is as follows.

---

[17] $\cdot$ makes non-terminating feature of justification logics more drastic, but proof search procedure in logics containing 4 may not terminate, either.

Claim 4.1.21 *Let* $\sigma_1\Phi_1$ *and* $\sigma_2\Phi_2$ *be prefixed signed formulas such that* $h_{12}(\sigma_1) \neq \sigma_2$ *and let* $\mathcal{B}_1^r$ *be the reduction of* $\mathcal{B}_1$ *using rule* (r) *on* $\sigma_1\Phi_1$ *and let* $\mathcal{B}_2^r$ *be the reduction of* $\mathcal{B}_2$ *using rule* (r) *on* $h_{12}(\sigma_1)\Phi_1$*. If rule* (r_2) *can be applied to* $\sigma_2\Phi_2$ *in* $\mathcal{B}_2$*, then it can be applied in* $\mathcal{B}_2^r$*.*

Proof of the claim: The cases of superposition can be reformulated along the same lines. Following Massacci, we present the only substantial case, namely, case 4.

For $\mathcal{B}_1$, we have the prefixed formulas $\sigma_1 F\Box\varphi$ and $\sigma_1 F\Box\psi$ and prefixes $\sigma_1.n_1$ and $\sigma_1.m_1$ are new. The same conditions (changing the subscript) hold for $\mathcal{B}_2$. We also have $h_{ij}(\sigma_i) = \sigma_j$. We have no constraints on $\sigma_i.n_i$ or $\sigma_j.n_j$ because they are new.

As in the proof for $x = x'$, renamings must be updated when the branch $\mathcal{B}_1$ reduces to $\mathcal{B} \cup \{\sigma_1.n_1 F\varphi, \sigma_1.m_1 F\psi\}$ and $\mathcal{B}_2$ is reduced to $\mathcal{B} \cup \{\sigma_2.n_2 F\varphi, \sigma_2.m_2 F\psi\}$.

$$h'_{ij}(s) = \begin{cases} h_{ij}(\sigma_i).n_j & \text{if } s = \sigma_i.n_i \\ h_{ij}(\sigma_i).m_j & \text{if } s = \sigma_i.m_i \\ h_{ij}(s) & \text{otherwise} \end{cases}$$

By hypothesis, $h_{ij}(\sigma_i) = \sigma_i$ and the new mappings h'_{ij} give the desired renamings. ⊠ (lemma 4.1.19)

4.1.4.4 Numerical variants of closed tableaux

Recall that one of the sources of the problem in translation is a case in which reverse rules are involved. In fact, the most difficult case is the one in which after reverse rules are applied we use the same prefixes as we already used (namely, we go back and forth) since this makes $\nu_{K,4}$ rules inevitably separated from the pertinent application of π-rule. (Note that reverse rules have to be applied between π and $\nu_{K,4}$.) However, it turns out that such back and forth steps are not necessary to close a tableau, although it facilitates closure of a tableau. This is because whenever we use a reverse rule, it suffices to start over with a new prefix. Here we show that this is always possible by using renaming of a prefix.

We prove that we can eliminate the case ii) 2) in 1.2, but to do that, we start proving a more general proposition stating the closure of a "numerical variant" of a closed tableau in a prefixed tableau system. By "a numerical variant of a prefixed formula," we mean a prefixed formula that is different only in the number of the prefix from the formula. This can be taken as constructing a case of Massacci's renaming function defined in the previous section. We call a prefix $h(\sigma)$ a numerical variant of σ if there is at least one i-th element of σ s.t. $h(n_i) \neq n_i$. As a corollary of proposition 4.1.14, we can prove the following.

Proposition 4.1.22 *If a tableau $\mathcal{T}_i$ is closed, then its numerical variant $\mathcal{T}_j$ is also closed, provided that it satisfies the preconditions of applications for*

π, ν, and E rules.

Proof Since the mapping h preserves all applications of rules by the above proposition 4.1.14, closure of a tableau is also preserved under the mapping h. $\boxtimes$

Now we apply the general idea of handling numerical variants of closed tableaux to the aforementioned combination of E4r (or EFr) and $\nu_{K,4}$. First, we introduce a terminology. We call a "reuse of prefix" an application of $\nu_{K,4}$ on a prefix that has already been used for applying reverse rules.[18] We prove that such applications of $\nu_{K,4}$ can be eliminated from a proof with such applications.

Lemma 4.1.23 *If a tableau is closed by using reuses of prefixes by $\nu_{K,4}$-rules, then a closed tableau can be constructed without using such applications.*

Proof A reuse of a prefix gives a reduction sequence as follows ($\sigma F\Box \in \mathcal{B}_1, \mathcal{B}_2 = \mathcal{B}_1 \cup \{\sigma.nFA\}$, $\sigma T\Box C \in \mathcal{B}_j$, $\mathcal{B}_{j+1} = \mathcal{B}_j \cup \{\sigma.nTC\}$. E4r or EFr is used somewhere between $\mathcal{B}_2$ and $\mathcal{B}_j$ on the branch. The last step reuses $\sigma.n$).

$$\mathcal{B}_1 \longrightarrow_{\pi} \mathcal{B}_2 \ldots \longrightarrow_{E4r(EFr)} \ldots \mathcal{B}_j \longrightarrow_{\nu_{K,4}} \mathcal{B}_{j+1} \ldots \mathcal{B}_k$$

---

[18] We use the word "reuse" only for cases in which reverse rules are applied. There are some other cases in which a prefix is "re-used" without reverse rules being applied. But these cases do not require any special treatment. Thus, there is no point of introducing a special terminology.

It may appear to be impossible to remove E4r or EFr between π and $\nu_{K,4}$ since $\nu_{K,4}$ is applicable only via E4r or EFr (See the example in the previous subsection.) But we claim that such a reuse of $\sigma.n$[19] for $\nu_{K,4}$-rule can be eliminated via the following procedure.

1) Pick up (on the same branch) the application of π-rule to a formula $\sigma F \Box \psi$ where $\sigma.n F \psi$ is obtained ($\sigma.n$ is produced) by the application of π rule.

2) Duplicate the $\sigma F \Box \psi$, and apply *exactly the same rules* (starting with π with $\sigma.m$ s.t. $m \neq n$) as applied in constructing the original tableau before the application of $\nu_{K,4}$, and then apply $\nu_{K,4}$ by using $\sigma.m$, etc.

3) After constructing a closed tableau with new prefixes, delete the relevant part of the original tableau that has the reused prefixes.

Applied this procedure, the reduction sequence will be modified as follows.

$$\mathcal{B}_1 \longrightarrow_{\pi_1} \mathcal{B}_2 \ldots \longrightarrow_{E4r(EFr)} \mathcal{B}_j \longrightarrow_{\pi_2} \mathcal{B}'_2 \ldots \mathcal{B}'_j \longrightarrow_{\nu_{K,4}} \mathcal{B}'_{j+1} \ldots \mathcal{B}'_k.$$

$\mathcal{B}_i = \mathcal{B}'_1, \mathcal{B}'_2 \ldots \mathcal{B}'_j$ are sets of formulas where each contains numerical variants of $\mathcal{B}_1, \ldots \mathcal{B}_j$ (the same rules as applied between $\mathcal{B}_1$ and $\mathcal{B}_j$ are applied. $\mathcal{B}_2 = \mathcal{B}_j \cup \{\sigma.mFA\}$, $\mathcal{B}'_{j+1} = (\mathcal{B}'_j \cup \{\sigma.mTC\}) \backslash \{\sigma.nTC\}$.)

E.g. The example of a tableau proof given above is modified as follows.

---

[19] A generalization to $\sigma.\sigma'$ is straightforward.

$1F\Box(t : \Box A \to A)$

$1.1Ft : \Box A \to A;$

$1.1Tt : \Box A\ ;\ 1.1FA$

$1Tt : \Box A;\ 1T\Box A$

$\mathbf{1.2Ft} : \Box\mathbf{A} \to \mathbf{A};$

$\mathbf{1.2Tt} : \Box\mathbf{A};\ \mathbf{1.2FA}$

$\mathbf{1Tt} : \Box\mathbf{A};\ \mathbf{1T}\Box\mathbf{A}$

$1.2T\Box A$

$1.2TA$

$\times$

This procedure can be taken to contain the process of explicitly constructing a renaming mapping h from the old tableau to the new tableau. Since a tableau is a set, we can duplicate the relevant part and take the numerical variant of the part. By proposition 21, if the original tableau is closed, then the new tableau is also closed. (Subtraction of the reuse of prefixed tableau does not affect the closure, since the numerical variant (on the same branch) is guaranteed to be closed by the proposition 21.) Note also that after this modification we do *not* have to apply the second π *before* the first applications of reverse rules (since a prefix $\sigma.n$ is already available to apply E4r or EFr). Then we can use another prefix $\sigma.m$ to apply $\nu_{K,4}$ (after the second π). By using this procedure, we can eliminate reuses of the prefix $\sigma.n$. We can apply this procedure for finitely many reuses of prefixes in a given tableau proof. Hence, we can eliminate all reuses of prefixes from it. ⊠

Note: Elimination of reuses of prefixes is not implied by proof confluence, since the manipulation is different from changing the order of applications rules. Also, the presence of reuses of prefixes does not spoil proof confluence, since it introduces no possibility of backtracking. (But it so happens that if we keep the prefixes in the reuses, then proof confluence is not flexible enough to make the prefixed tableau translatable into a hypersequent proof.)

4.1.4.5 Prefixed tableau proofs translatable to hypersequent proofs

Having these preparations, we are ready to transform any given prefixed tableau proof into a prefixed tableau proof translatable to a hypersequent proof. To make mathematically precise the idea described in 4.1.4.2, we first define a normal form of a proof in the prefixed tableau system and prove a proposition.

Definition 4.1.24 *We call a prefixed tableau proof in a* π-$\nu_{K,4}$ *normal form when for any application of* $\nu_{K,4}$*, there exists an application of* π *such that the prefix used in the application of* $\nu_{K,4}$ *is introduced by the application of* π *and there exists no application of rule* x *such that* $\pi \gtrdot x \gtrdot \nu_{K,4}$ *(except other applications of* $\nu_{K,4}$*).*

Proposition 4.1.25 *Suppose there exists a tableau proof of* $\{1F\varphi\}$ *in the prefixed tableau systems* **TS4LPN**. *Then, this tableau proof can be effectively transformed into a tableau proof that is in a* π-$\nu_{K,4}$ *normal form.*

Proof We have two cases, i.e., a case in which prefixes are reused and a case in which they are not reused.

Case 1. **The case in which prefixes are reused.**

By lemma 4.1.22, we can eliminate all reuses of prefixs $\sigma.n$, preserving closure. Therefore, the first case is reducible to the second.

Case 2. **The case in which prefixes are not reused**

We give a proof by induction on the number of other rules occurring between a $\nu_{K,4}$-rule and the application of π-rule that introduced the prefix the application of $\nu_{K,4}$ rule uses. Since we have eliminated reuses of prefixes, the only case with a fixed order of applications of rules based on prefixes is π-$\nu_{K,4}$. For other cases of rules, applications of rules can be reordered under some natural constraints (described below). By the novelty of $\sigma.n$ in the application of π, some formulas cannot occur before the application of π. In all the subcases, a formula occurring between the pertinent π and $\nu_{K,4}$ can be moved out as follows (here $\sigma.n$ is the prefix obtained by the application of π):

Case (A) the prefix σ' on the premise or the conclusion is $\sigma' \geq \sigma.n$ or the premise of the tableau rule derives from a formula with the prefix $\sigma' \geq \sigma.n$;

Case (B) Otherwise.

For (A), we move formulas occurring between π and $\nu_{K,4}$ *after* $\nu_{K,4}$. For (B), we can move these *before* the application of π.

For any application of a rule, we can apply the idea of proof confluence and we can move them before π or after $\nu_{K,4}$, depending on which condition ((A) or (B)) is satisfied by the prefix. Suppose the number of applications is n. Pick one application after π. This can be any of all the rules in TS4LPN except $\nu_{K,4}$. We have many cases, but for each case, depending on whether a particular case satisfies the condition A or B, each of these cases can be moved out of the interval before the application of π or after the application of $\nu_{K,4}$. This claim can be simply guaranteed by proof confluence, which is a corollary of lemma 4.1.18. (So, we omit checking correctness of the claim case by case.) Note that proof confluence allows shuffling of applications of rules except $\pi \gtrdot \nu_{K,4}$ since reuses of formulas are already eliminated (see footnote 15 for other exceptions that does not affect the argument). Hence, these cases cover all the cases. Therefore, for one step, all these rules can be moved out. Then the number of the rules applied between π and $\nu_{K,4}$ can be reduced by 1. By IH, for any number strictly less than n, the statement holds. Therefore, all n applications of rules can be removed between the application of π and $\nu_{K,4}$. ⊠

E.g. The example given in the previous section is not yet in a π-$\nu_{K,4}$ normal form. By appealing to proof confluence, the tableau rules can be rearranged and modified into a π-$\nu_{K,4}$ normal form as follows.

$1F\Box(t : \Box A \to A)$

$1.1.Ft : \Box A \to A$

$1.1.Tt : \Box A$; $1.1.FA$

$1Tt : \Box A$

$1T\Box A$

$\mathbf{1.2.Ft : \Box A \to A}$;

$1.2T\Box A$; $1.2TA$

$\mathbf{1.2.Tt : \Box A}$; $\mathbf{1.2.FA}$

×

Once we obtain a π-$\nu_{K,4}$ normal form, we replace all the applications of π by $\pi^\sharp$.

4.1.4.6 Inductive proof of preserving provability under the translation

We introduced an auxiliary prefixed tableau system $\mathsf{TS4LPN}^\sharp$. In this system, instead of using π-rule, we use $\pi^\sharp$-rule so that we can handle π-rule and $\nu_{K,4}$ simultaneously.

Prefixed tableau system $\mathsf{TS4LPN}^\sharp$ works well for handling some discrepancies between prefixed tableau systems and hypersequent calculi, but it apparently makes complicated the induction on the depth of a tableau proof. Handling $\nu_{K,4}$ produces complication. To make the inductive proof simpler, we use a subsystem $\mathsf{TS4LPN}^{\sharp\circ}$ in which we keep $\pi^\sharp$ rule, but we remove $\nu_{K,4}$. We treat $\nu_{K,4}$ as admissible rules (not primitive rules) in $\mathsf{TS4LPN}^{\sharp\circ}$. Our translation from $\mathsf{TS4LPN}^{\sharp\circ}$ to the hypersequent calculus $\mathsf{HS4LPN}^-$ is shown to preserve provability.

We claim that due to the theorem stating the existence of π-$\nu_{K,4}$ normal form, $\nu_{K,4}$ can be absorbed into $\pi^\sharp$-rule.

Proposition 4.1.26 *1. In* $\mathsf{TS4LPN}$*,* $\pi^\sharp$ *is a derived rule. 2. For any* φ*, if* φ *has a tableau proof in* $\mathsf{TS4LPN}$*, then* φ *has a tableau proof in* $\mathsf{TS4LPN}^{\sharp\circ}$*.*

Proof 1. Consider any application of $\pi^\sharp$. An application of $\pi^\sharp$ can be clearly simulated by first applying π rule and then apply all possible applications of $\nu_{K,4}$ in the given premise.

2. By the above observation, we start from $\mathsf{TS4LPN}$. A proof should be given by induction on the depth of tableau proof of $\mathsf{TS4LPN}$ in π-$\nu_{K,4}$ normal form. However, since obviously $\mathsf{TS4LPN}^{\sharp\circ}$ can have all the rules of $\mathsf{TS4LPN}^\sharp$ except $\nu_{K,4}$ rules, it suffices to prove that $\pi^\sharp$ rule can cover all the applications of $\nu_{K,4}$. First, we replace all the applications of π-rule in π-$\nu_{K,4}$ normal form by $\pi^\sharp$-rule. Due to the π-$\nu_{K,4}$ normal form theorem, any application of $\nu_{K,4}$ must immediately follow an application $\pi^\sharp$. (Note that it suffices if we have the preconditions of the applications of $\nu_{K,4}$ that are actually used in the original proof in $\mathsf{TS4LPN}$.) This can be taken as a proof in $\mathsf{TS4LPN}^\sharp$. In this form of proof in $\mathsf{TS4LPN}^\sharp$, all the applications of these rules have the following form. Let $\sigma F\Box A \in S$ and $\sigma T\Box B_1, \ldots, \sigma T\Box B_m \in S$ (list up all the formulas of the form $\sigma T\Box B_i$ in S.)

$\sigma F\Box A$

$\sigma.nFA$, $\sigma.nT\Box B_1$, $\sigma.nTB_1, \ldots, \sigma.nT\Box B_m, \sigma.nTB_m$ (call this $S^\sharp_{\sigma.n}$) by $\pi^\sharp$

$\sigma.nT\Box B_j$, $\sigma.nTB_j$

$\ldots$

$\sigma.nT\Box B_k$, $\sigma.nTB_k$ (by several applications of $\nu_{K,4}$. $1 \leq j \leq k \leq m$.)

Clearly, $\sigma.n\Box TB_j, \sigma.nTB_j, \ldots, \sigma.nT\Box B_k, \sigma.nTB_k \subseteq S^\sharp_{\sigma.n}$ (due to the precondition of application of $\nu_{K,4}$, these formulas must be a subset of $S^\sharp_{\sigma.n}$). Hence, anything that can be derived by using a sequence of applications of rules $\pi^\sharp$, $\nu^1_{K,4}, \ldots, \nu^k_{K,4}$ (up to m) can already be derived in one-step application of $\pi^\sharp$. Therefore, $\mathsf{TS4LPN}^{\sharp\circ}$ is sufficient to derive all the formulas that $\mathsf{TS4LPN}^\sharp$ can derive and, hence, $\nu_{K,4}$ are dispensable (hence, admissible in $\mathsf{TS4LPN}^{\sharp\circ}$). ⊠

Theorem 4.1.27 *Let S be a finite set of prefixed signed formulas. If there is a closed tableau for S using the prefixed tableau for $\mathsf{S4LPN}$, then the hypersequent S^s is provable in $\mathsf{HS4LPN}^-$.*

Proof Suppose that there is a prefixed tableau proof in $\mathsf{TS4LPN}$ of S. Fix one. We call it $\mathcal{T}_1$. By using the manipulations of prefixed tableau proofs given in the previous section, we transform $\mathcal{T}_1$ into a prefixed tableau proof in π-$\nu_{K,4}$ normal form. Call the proof $\mathcal{T}_1^*$. (It must be clear that the normal form is not unique.)

Modify $\mathcal{T}_1^*$ into $(\mathcal{T}_1^*)^\sharp$ by replacing all applications of π by $\pi^\sharp$ (given preconditions). (We do not have to apply $\nu_{K,4}$ in $\mathcal{T}_1^*$, but this replacement amounts to applying all possible cases of $\nu_{K,4}$. This introduces some redundancy in this translation, but it preserves derivability and introduces no harm, since obviously these applications are derivable in the original system, too. Since a tableau is a set of formulas, this is the only difference between $\mathcal{T}_1^*$ and $(\mathcal{T}_1^*)^\sharp$).) We can take the resulting proof as a proof in $\mathsf{TS4LPN}^{\sharp\circ}$ due to the observation in the proposition 4.1.25.

Then apply the mapping s from the prefixed tableau $(\mathcal{T}_1^*)^\sharp$ (= a set of prefixed formulas) to a hypersequent $((\mathcal{T}_1^*)^\sharp)^s$. (The mapping has nothing to do with inference rules. It is a mapping from the deductive meta-language of $\mathsf{TS4LPN}$ to the deductive meta-language of $\mathsf{HS4LPN}$).

Now we show by induction on the depth d of a prefixed tableau in $\mathsf{TS4LPN}^{\sharp\circ}$ that for any $(\mathcal{T}_1^*)^\sharp$, $((\mathcal{T}_1^*)^\sharp)^s$ is provable in $\mathsf{HS4LPN}^-$.

Here the depth is a number $d \geq 0$ such that there is a closed prefixed tableau in $\mathsf{TS4LPN}$ for S with d applications of tableau rules. Suppose a tableau for S closes with depth d and, by IH, the theorem holds for sets that close with the depth less than d. (This means that IH makes sense only for $d \geq 1$ since we take $d-1$ for IH. $d = 0$ is the base case, where S is already a closed tableau.) Suppose we have made the first application of a tableau rule. Then, to obtain a closed tableau, we only need $d-1$ applications of the rules. For any tableau rule, after this first application we can apply IH. Then, there must be a closed tableau in $\mathsf{TS4LPN}^{\sharp\circ}$ for a set of formulas obtained as the result of applying one of the rules, and such a closed tableau must be mapped into a hypersequent by the mapping s (such a hypersequent must be provable in $\mathsf{HS4LN}^-$ since by the definition of closure of a tableau and the mapping s, the hypersequent that is the image of a closed tableau can be taken as a result of applying IW and EW to an axiom in $\mathsf{HS4LPN}$). Hence, the remaining step is to prove in $\mathsf{HS4LPN}^-$, for each rule of $\mathsf{TS4LPN}$, the hypersequent obtained by taking the image of the mapping s of the set of the prefixed signed formulas that we had before we apply the rule. We omit α, β cases, since they are straightforward.

(**Note**: In the following proof, some lines look redundant since hypersequents are sets and we do not have (internal or external) contraction rules. However, we put these lines to make it more perspicuous how derivations work. In these cases, we put (set) on the right side of the step. Also, several applications of the same rule, which must be sufficiently clear from the context, are suppressed without using double lines.)

Case 1 ET-rule :

Suppose $\sigma Tt : \varphi \in S$. By the assumption, such an S has a closed tableau with depth d. Since the first application of a rule in this tableau is ET-rule, $S \cup \{\sigma T\varphi\}$ has a closed tableau with depth $d - 1$. By IH, $(S \cup \{\sigma T\varphi\})^s$ provable in the hypersequent calculus $\mathsf{HS4LPN}^-$. By definition of s, the image of the set given above by s has the form $\Gamma_1 \Rightarrow \Delta_1 | \dots | \Gamma_\sigma, t : \varphi, \varphi \Rightarrow \Delta_\sigma | \dots | \Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n} | \dots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$. We show below that S^s, which is equal to $\Gamma_1 \Rightarrow \Delta_1 | \dots | \Gamma_\sigma, t : \varphi \Rightarrow \Delta_\sigma | \Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n} | \dots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$, is provable in $\mathsf{HS4LPN}^-$.

$$\dfrac{\dfrac{\Gamma_{\sigma_1} \Rightarrow \Delta_{\sigma_1}|\dots|\Gamma_\sigma, t:\varphi, \varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}{\Gamma_{\sigma_1} \Rightarrow \Delta_{\sigma_1}|\dots\Gamma_\sigma, t:\varphi, t:\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n}, \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\,\text{Lt}}{\Gamma_{\sigma_1} \Rightarrow \Delta_{\sigma_1}|\dots|\Gamma_\sigma, t:\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\,\text{(set)}$$

Case 2. EK-rule :

Suppose $\sigma T t : \varphi \in S$. By the assumption, such an S has a closed tableau with depth d. Since the first application of a rule in this tableau is EK-rule, $S \cup \{\sigma.n T \varphi\}$ has a closed tableau with depth $d-1$. By IH, $(S \cup \{\sigma.n T\varphi\})^s$ is provable in $\mathsf{HS4LPN}^-$. By definition of s, S^s is as follows: $\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma, t : \varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$. This can be proven by the following derivation.

$$\dfrac{\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma, t:\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n}, \varphi \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma, t:\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n}, t:\varphi \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\,Lt}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma, t:\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|t:\varphi \Rightarrow |\dots|\Gamma_{\sigma.m} \Rightarrow \Delta_{\sigma.m}}}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma, t:\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}$$

(The second step is Splitting, and the third uses IW and (set).)

Case 3. $E4$-rule :

Suppose $\sigma T t : \varphi \in S$. By the assumption, such an S has a closed tableau with depth d. Since the first application of a rule in this tableau is $E4$-rule, $S \cup \{\sigma.n T t : \varphi\}$ has a closed tableau with depth $d-1$. By IH, $(S \cup \{\sigma.n T t : \varphi\})^s$ is provable in $\mathsf{HS4LPN}^-$. By definition of s, S^s is as follows: $\Gamma_1 \Rightarrow \Delta_1|\dots|t : \varphi, \Gamma_\sigma \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$. This can be proven by the following derivation.

$$\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1|\dots|t:\varphi, \Gamma_\sigma \Rightarrow \Delta_\sigma|\dots|t:\varphi, \Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}{\Gamma_1 \Rightarrow \Delta_1|\dots|t:\varphi, \Gamma_\sigma \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|t:\varphi \Rightarrow |\dots|\Gamma_{\sigma.m} \Rightarrow \Delta_{\sigma.m}}}{\Gamma_1 \Rightarrow \Delta_1|\dots|t:\varphi, \Gamma_\sigma \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m}, \Rightarrow \Delta_{\sigma_m}}$$

(The first step is Splitting, and the second uses IW and (set).)

Case 4. $E4r$-rule :

Suppose $\sigma.nTt : \varphi \in S$. By the assumption, such an S has a closed tableau. Also, since the first application of a rule in the tableau is E4r, $S \cup \{\sigma Tt{:}\varphi\}$ has a closed tableau with depth $d-1$. By IH, $(S \cup \{\sigma Tt{:}\varphi\})^s$ is provable in the hypersequent calculus $\mathsf{HS4LPN}^-$, where this has the form $\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma, t{:}\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n}, t{:}\varphi \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$. We show S^s, which is equal to $\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma|\Gamma_{\sigma.n}, t : \varphi \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$, is provable only by using rules of the hypersequent calculus.

$$\dfrac{\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma, t{:}\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n}, t{:}\varphi \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma|t{:}\varphi \Rightarrow |\dots|\Gamma_{\sigma.n}, t{:}\varphi \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n}, t{:}\varphi \Rightarrow \Delta_{\sigma.n}|\Gamma_{\sigma.n}, t{:}\varphi \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma.n}, t{:}\varphi \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}$$

(The second step is Splitting, the third is IW, and the last step is (set).)

Case 5. EF-rule :

Suppose $\sigma Ft{:}\varphi \in S$. By the assumption, such an S has a closed tableau, and the first application of a rule in this case is EF. So, $S \cup \{\sigma.nFt{:}\varphi\}$ has a closed tableau of depth $d-1$. By IH, $(S \cup \{\sigma.nFt{:}\varphi\})^s$ is provable in the hypersequent calculus $\mathsf{HS4LPN}^-$, where this has the form $\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t{:}\varphi|\dots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}, t{:}\varphi|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$.

We show that S^s, which is equal to $\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t{:}\varphi|\dots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$, is provable only by using rules of the hypersequent calculus.

$$\begin{array}{c}
\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t{:}\varphi|\ldots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}, t{:}\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} \\ \hline
\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t{:}\varphi|\ldots| \Rightarrow t{:}\varphi|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} \\ \hline
\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t{:}\varphi|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t{:}\varphi|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} \\ \hline
\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t{:}\varphi|\ldots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}
\end{array}$$

(The second step is Splitting, the third is IW, and the last step is (set).)

Case 6. EFr-rule:

Suppose $\sigma.nFt : \varphi \in S$. By the assumption, such an S has a closed tableau. Also, since the first application of a rule in the tableau is EFr, $S \cup \{\sigma Ft{:}\varphi\}$ has a closed tableau with depth $d-1$. By IH, $(S \cup \{\sigma Ft{:}\varphi\})^s$ is provable in $\mathsf{HS4LPN}^-$, where this has the form $\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t : \varphi|\ldots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}, t : \varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$. We show S^s, which is equal to $\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}, t : \varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$, is provable only by using rules of $\mathsf{HS4LPN}^-$.

$$\begin{array}{c}
\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t{:}\varphi|\ldots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}, t{:}\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} \\ \hline
\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma| \Rightarrow t{:}\varphi|\ldots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}, t{:}\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} \\ \hline
\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma|\ldots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}, t{:}\varphi|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}, t{:}\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} \\ \hline
\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma|\ldots|\Gamma_{\sigma.n} \Rightarrow \Delta_{\sigma.n}, t{:}\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}
\end{array}$$

(The second step is Splitting, the third is IW, and the last step is (set).)

Case 7. Operational rules :

Subcase 7.1. !-rule :

Suppose $\sigma F!t : t : \varphi \in S$. By the assumption, such an S has a closed tableau. Also, the first application of a rule in the tableau is !-rule, and $S \cup \{\sigma Ft{:}\varphi\}$ has a closed tableau with depth $d-1$. By IH, $(S \cup \{\sigma Ft{:}\varphi\})^s$ is provable in $\mathsf{HS4LPN}^-$, i.e. $\mathsf{HS4LPN}^- \vdash \Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t : \varphi, !t : t : \varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$. The following proof is enough to show the provability of S^s in $\mathsf{HS4LPN}^-$.

$$\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t\!:\!\varphi, !t\!:\!t\!:\!\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, !t\!:\!t\!:\!\varphi, !t\!:\!t\!:\!\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\ \text{!-rule}}{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, !t\!:\!t\!:\!\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\ \text{(set)}$$

Subcase 7.2. +-rule :

Suppose $\sigma Ft + s : \varphi \in S$. By the assumption, such an S has a closed tableau. Also, the first application of a rule in the tableau is +-rule and $S \cup \{\sigma Ft\!:\!\varphi\}$ has a closed tableau with depth $d-1$. By IH, $(S \cup \{\sigma Ft\!:\!\varphi\})^s$ is provable in $\mathsf{HS4LPN}^-$, i. e. $\mathsf{HS4LPN}^- \vdash \Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t\!:\!\varphi, t+s\!:\!\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$. Then we can have the following proof, which shows the provability of S^s in $\mathsf{HS4LPN}$. ($s + t\!:\!\varphi$ case is similar.)

$$\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t\!:\!\varphi, t+s\!:\!\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t+s\!:\!\varphi, t+s\!:\!\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\ \text{+-rule}}{\Gamma_{\sigma_1} \Rightarrow \Delta_{\sigma_1}|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t+s\!:\!\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\ \text{(set)}$$

Subcase 7.3. $\cdot$-rule:

Suppose $\sigma t{\cdot}s\!:\!\varphi \in S$. By the assumption, such a S has a closed tableau. In particular, in this case, the rule produces two branches, so $S \cup \{\sigma Ft\!:\!\psi \to \varphi\}$ and $S \cup \{\sigma Fs : \psi\}$ have closed tableaus with depth $d-1$. By IH, then, $(S \cup \{\sigma Ft : \psi \to \varphi\})^s$ and $(S \cup \{\sigma Fs : \psi\})^s$ are provable in $\mathsf{HS4LPN}^-$, i.e., $\mathsf{HS4LPN}^- \vdash \Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t : \psi \to \varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$ and $\mathsf{HS4LPN}^- \vdash \Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, s\!:\!\psi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$. Then we have the following proof, which shows the provability of S^s in $\mathsf{HS4LPN}^-$. (The first line is an application of $\cdot$-rule. Note that this is a two-premise rule)

$$\dfrac{\dfrac{\begin{array}{c}\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t\!:\!\psi \to \varphi, t{\cdot}s\!:\!\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}\\ \Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, s\!:\!\psi, t{\cdot}s\!:\!\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}\end{array}}{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t{\cdot}s\!:\!\varphi, t{\cdot}s\!:\!\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}}{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, t{\cdot}s\!:\!\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\ \text{(set)}$$

Case 8. ν_T-rule.

Suppose $\sigma T\Box\varphi \in S$. By the assumption, such an S has a closed tableau with depth d. Since the first application of a rule in this tableau is ν_T-rule, $S \cup \{\sigma T\varphi\}$ has a closed tableau with depth $d-1$. By IH, $(S\cup\{\sigma T\varphi\})^s$ is provable in $\mathsf{HS4LPN}^-$. By definition of s, S^s is as follows: $\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma, \Box\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$. Then we have the following derivation, which shows provability of S^s in $\mathsf{HS4LPN}^-$.

$$\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma,\Box\varphi,\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma_m}\Rightarrow\Delta_{\sigma_m}}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma,\Box\varphi,\Box\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma_m}\Rightarrow\Delta_{\sigma_m}}\,\mathrm{L}\Box}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma,\Box\varphi \Rightarrow \Delta_\sigma|\dots|\Gamma_{\sigma_m}\Rightarrow\Delta_{\sigma_m}}\,(\mathrm{set})$$

Case 9. $\pi^\sharp$-rules

$\sigma F\Box\varphi \in S$, and $S^\sharp_{\sigma.n} = \{\sigma.nT\Box\psi, \sigma.nT\psi|\sigma\Box\psi \in S\}$. So, $(S^\sharp_{\sigma.n})^s = \Gamma^\sharp_{\sigma.n}$. Note that $\{\Box\psi|\Box\psi \in \Gamma^\sharp_{\sigma.n}\} \subseteq \Gamma_\sigma$. Then by IH, $(S \cup \{\sigma.nF\varphi\} \cup S^\sharp_{\sigma.n})^s$, which is identical to $\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\Gamma^\sharp_{\sigma.n} \Rightarrow \varphi|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$, is provable in $\mathsf{HS4LPN}^-$. We can derive S^s from the above hypersequent in $\mathsf{HS4LPN}^-$ as follows.

$$\begin{array}{c}
\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma,\Box\varphi|\Gamma^\sharp_{\sigma.n} \Rightarrow \varphi|\dots|\Gamma_{\sigma_m}\Rightarrow\Delta_{\sigma_m}\\ \hline
\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma,\Box\varphi|\Box\psi_1,\psi_1,\dots,\Box\psi_k,\psi_k \Rightarrow \varphi|\dots|\Gamma_{\sigma_m}\Rightarrow\Delta_{\sigma_m}\\ \hline
\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma,\Box\varphi|\Box\psi_1,\Box\psi_1,\dots,\Box\psi_k,\Box\psi_k \Rightarrow \varphi|\dots|\Gamma_{\sigma_m}\Rightarrow\Delta_{\sigma_m}\\ \hline
\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma,\Box\varphi|\Box\psi_1,\dots,\Box\psi_k \Rightarrow \varphi|\dots|\Gamma_{\sigma_m}\Rightarrow\Delta_{\sigma_m}\\ \hline
\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma,\Box\varphi|\Box\psi_1,\dots,\Box\psi_k \Rightarrow \Box\varphi|\dots|\Gamma_{\sigma_m}\Rightarrow\Delta_{\sigma_m}\\ \hline
\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma,\Box\varphi|\Gamma_\sigma \Rightarrow \Delta_\sigma,\Box\varphi|\dots|\Gamma_{\sigma_m}\Rightarrow\Delta_{\sigma_m}\\ \hline
\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma,\Box\varphi|\dots|\Gamma_{\sigma_m}\Rightarrow\Delta_{\sigma_m}
\end{array}$$

(The first step is the definition of $\Gamma^\sharp_{\sigma.n}$, the second is L$\Box$, the third is (set), the fourth is R$\Box$, the fifth is IW, and the last one is (set).)

Note: Due to the proposition 4.1.26, we can work in $\mathsf{TS4LPN}^{\sharp\circ}$. Therefore, we do *not* have $\nu_{K,4}$ rules as primitive rules in $\mathsf{TS4LPN}^{\sharp\circ}$.

Theorem 4.1.28 (Completeness) *If a* formula φ *is valid in the semantics for S4LPN, then HS4LPN⁻* $\vdash\Rightarrow\varphi$*.*

Proof Use completeness of the prefixed tableau system. Then, apply Lemma 4. This is a special case of the lemma where $S = \{1F\varphi\}$. ⊠

As a corollary, we can semantically prove cut-admissibility for HS4LPN.

Corollary 4.1.29 *If HS4LPN* $\vdash\Rightarrow\varphi$*, then HS4LPN⁻* $\vdash\Rightarrow\varphi$

Proof By theorem 4.1.12, HS4LPN is sound with respect to the Hilbert-style system, and the Hilbert-style system is sound with respect to the Kripke semantics. By the above theorem, $\mathsf{HS4LPN}^{-}$ is complete with respect to the Kripke semantics. Hence, the statement follows. ⊠

Here are applications of cut-admissibility. We first state a conservativeness result.

Proposition 4.1.30 *For any formula* φ *that does not contain any proof-polynomial, if S4LPN* $\vdash\varphi$*, then S4* $\vdash\varphi$*.*

Proof Although the subformula property does not hold for HS4LPN, the hypersequent calculus has a feature that once we introduce a proof term, the proof term would never disappear from the cut-free proof in which it is introduced. This is obvious from the forms of each operational rule and cut-admissibility of the system. So, if you can prove a statement that is stated in purely modal language, then there must exist a cut-free proof of the statement that does not contain any formula that has a proof term in it. Naturally, all the applications of the rules in the proof must come from S4 part of the system. Therefore, the statement is provable in S4. ⊠

This conservativeness result suggests a possibility of obtaining some version of the disjunction property. It is well-known that the modal disjunction property in the following form holds for S4.

(MDP) $\mathsf{S4} \vdash \Box\varphi \vee \Box\psi \Longrightarrow \mathsf{S4} \vdash \varphi$ or $\mathsf{S4} \vdash \psi$.

It is a natural question whether some analogue of MDP holds for our S4LPN. MDP does not hold for the entire system. However, if we confine ourselves to a particular context, we can obtain some refined version of MDP. In particular, due to conservativeness of modal language and LP language, as far as a theorem of S4LPN contains only modal operators but does not contain any proof polynomial, we can have the disjunction property for both S4 and LP. In other words, S4LPN combines two logics without interfering either fragment.

Proposition 4.1.31 *For any formula φ, ψ that does not contain any proof-polynomial, if S4LPN $\vdash \Box\varphi \vee \Box\psi$, then S4LPN $\vdash \varphi$ or S4LPN $\vdash \psi$.*

Proof Due to the fact that $\Box\varphi \vee \Box\psi$ does not contain any proof-term, we can apply the above proposition 4.1.30 to this statement. So, $\mathsf{S4} \vdash \Box\varphi \vee \Box\psi$. Apply the disjunction property for S4. Then, $\mathsf{S4} \vdash \varphi$ or $\mathsf{S4} \vdash \psi$. So, $\mathsf{S4LPN} \vdash \varphi$ or $\mathsf{S4LPN} \vdash \psi$. ⊠

The opposite combination of the problem of conservativeness, namely whether S4LPN is a conservative extension of LP, may not be so simple, since in this case we may lose some modal formula due to the rule $R\cdot$. Cut-admissibility does not guarantee that S4LPN is a conservative extension, but LP has another version of the disjunction property due to its constructive feature. The following is a corollary of the cut-elimination theorem for LP (it is also proven semantically in [96]).

For any $t{:}\varphi$ and $s{:}\psi$, $\mathsf{LP}\vdash t{:}\varphi \vee s{:}\psi \Longrightarrow \mathsf{LP}\vdash t{:}\varphi$ or $\mathsf{LP}\vdash s{:}\psi$.

A similar statement with respect to the language of S4LPN holds for S4LPN.

Proposition 4.1.32 *For any $t{:}\varphi$ and $s{:}\psi$, $\mathsf{S4LPN} \vdash t{:}\varphi \vee s{:}\psi \Longrightarrow \mathsf{S4LPN} \vdash t{:}\varphi$ or $\mathsf{S4LPN} \vdash s{:}\psi$.*

Proof (Proof sketch) If we take a cut-free proof, then the last application of a rule is R∨ (since our hypersequents are sets). Hence, the line before the last has the form $\Rightarrow t : \varphi, s : \psi$. We have only two rules to be applied in order to obtain this sequent, i.e., either IW or a rule for a proof-term, although we have two (times two) patterns of applications of IW, either from $\Rightarrow t : \varphi$ or from $\Rightarrow t : \varphi, s : \psi | \Rightarrow t : \varphi$. At least one of them $t : \varphi$, $s : \psi$ is introduced by a rule for a proof-term.

Since the outermost logical symbol of each formula ("o.l.s") is a proof-term, the premise(s) of a rule for a proof-term must have a (simpler) proof-term as the o.l.s. Hence, for every step upwards, the only possible application of a rule is either IW, EW, or a rule for a proof-term until we get a variable or a constant. If IW or EW whose premise has only one formula is used at some point, then we go to the premise of IW or EW, and reconstruct a proof of $\Rightarrow t : \varphi$ or $\Rightarrow s : \psi$. Otherwise, on each branch of a proof-tree we reach a line that has formulas with simple proof terms. Then, at least one of the o.l.s of the formulas is a proof constant and satisfies the condition of a constant specification. Otherwise, there is no rule to apply upwards. (Except IW or EW. But in some further step we have the same problem since not all formulas can be introduced by IW or EW.) This contradicts the provability of $\Rightarrow t : \varphi, s : \psi$. Since a constant specification in our system is just one formula on the succedent, one of them must be introduced by IW or EW. Either way, we have traced to the initial line of a proof of $\Rightarrow t : \varphi$ or $\Rightarrow s : \psi$.
⊠

4.1.5 Prefixed Tableaux and Hypersequents for S4LP

The prefixed tableau approach has an advantage in covering a wider range of modal logics than destructive ones can cover, due to its flexibility in prefixed tableau systems. A prefixed tableau system has been particularly helpful in formulating the mixed negative introspection $\neg t : \varphi \to \Box\neg t : \varphi$, since it is difficult to formulate a destructive tableau rule corresponding to this axiom. However, note that it was S4LP, not S4LPN, which was formulated first as a system combining LP and S4 ([58]). So, it is a natural question whether we can formulate a prefixed tableau system for S4LP and a hypersequent calculus for S4LP. In the Hilbert-style system, S4LP is the subsystem of S4LPN obtained by omitting the axiom $\neg t{:}\varphi \to \Box\neg t{:}\varphi$. It is indeed possible to formulate a prefixed tableau system for S4LP. In the following, we present both TS4LP and HS4LP. However, it turns out that there is more than one way in formulating S4LP in a prefixed tableau system. We officially adopt one of the two different ways of formulating the system in order to make it possible to translate prefixed tableau proofs to hypersequent proofs. With some minor twist, our method of proving cut-admissibility of a hypersequent calculus via a translation from a prefixed tableau system can also be applied to S4LP.

Before going into definitions, let us note that S4LP in Hilbert-style axiomatic formulation is complete with respect to two different semantics. The one is a special case of the other one. In a general approach, we use a semantics that has two accessibility relations R and R^e and requires $R \subseteq R^e$. (This distinction is introduced by Artemov and Nogina [8].) The other case is the one where we have $R = R^e$. Here we consider only the latter one.

A prefixed tableau system for S4LP is defined by omitting EFr and E4r from TS4LPN. Explicit rules of the system are the following:

$$\text{EK}\ \frac{\sigma Tt{:}\varphi}{\sigma.nT\varphi}\ (\sigma.n \text{ is used.}) \qquad \text{ET}\ \frac{\sigma Tt{:}\varphi}{\sigma T\varphi} \qquad \text{E4}\ \frac{\sigma Tt{:}\varphi}{\sigma.nTt{:}\varphi}\ (\sigma.n \text{ is used.})$$

Rules for Classical Propositional Logic, Operational Rules on a signed formula of the form $Ft : \varphi$, and Modal Rules are are the same as S4LPN. Constant Specification Rules are restricted to S4LP axioms. Fitting-style Kripke semantics for S4LP is obtained by modifying the Kripke model for S4LPN as follows: instead of taking R^e to be an equivalence relation on K s.t. $R \subseteq R^e$, we take $R = R^e$. Here R is a reflexive and transitive relation on K. Accordingly, our interpretation of R^e for S4LP becomes as follows: $\sigma \leq \sigma' \Longrightarrow \mathcal{N}(\sigma)R^e\mathcal{N}(\sigma')$.

Lemma 4.1.33 *Suppose $\mathcal{T}$ is a satisfiable tableau of S4LP. If any tableau rule for S4LP is applied to $\mathcal{T}$, then the resulting tableau is still satisfiable.*

Proof The proof is similar to that of S4LPN. In particular, we can modify the proof of the relevant cases (EK, ET, E4) of S4LPN by replacing $1 \leq \sigma$ and $1 \leq \sigma'$ with $\sigma \leq \sigma.n$ (EK, E4) or $\sigma \leq \sigma$ (ET). The proof obviously goes through. ⊠

Now we can state and prove the soundness theorem.

Theorem 4.1.34 (Soundness) *If φ has a prefixed S4LP-tableau proof, then φ is valid in all models of S4LP.*

Proof Similar to the case of S4LPN. ⊠

Concerning completeness, there is no essential difference in Lindenbaum-Henkin construction in the case of S4LPN and its subsystems, in particular S4LP, except the use of S4LP-consistency when you construct a maximal consistent set. We construct a canonical Kripke model $\mathcal{K} = (K, R, R^e, \mathcal{E}, \mathcal{V})$ for S4LP via Lindenbaum-Henkin construction. However, we have two modifications compared with the case of S4LPN. (1) we have $R = R^e$ (R^e is just a reflexive and transitive relation), and (2) in order for the evidence function to satisfy monotonicity, we need to modify the definition of $\mathcal{E}$. R^e for S4LP is defined as follows:

1. $\sigma R^e \sigma'$ iff σ is an initial segment of σ' ($\sigma \leq \sigma'$).

The model satisfies the condition $R = R^e$, since new prefixes are introduced only via π-formulas in the construction. R^e so defined obviously satisfies reflexivity and transitivity. Due to this identity, we use the symbol R (instead of R^e) in the following argument to simplify the notation. The evidence function $\mathcal{E}$ is defined as follows.[20]

2. $\varphi \in \mathcal{E}(\sigma, t)$ iff, for any $\sigma' \in K$, $\sigma \leq \sigma' \Rightarrow \sigma' F t{:}\varphi \notin S_\omega$.

---

[20] This definition is due to Evan Goris (personal communication), although his version is defined on destructive tableau system.

Now we give a proof that the conditions of evidence function for S4LP are satisfied.

Proposition 4.1.35 *The evidence function defined above satisfies the following conditions: (A) monotonicity, (B) closure conditions, (C) constant specification.*

Proof (A) (Monotonicity) We show $\sigma R\sigma'$ and $\varphi \in \mathcal{E}(\sigma, t)$ implies $\varphi \in \mathcal{E}(\sigma', t)$.

Suppose (1) $\sigma R\sigma'$, (2) $\varphi \in \mathcal{E}(\sigma, t)$ and (3) $\sigma' R\sigma''$. By (2), $\sigma R\sigma'' \Rightarrow \sigma'' Ft{:}\varphi \notin S_\omega$. By (1), (2) and transitivity, $\sigma R\sigma''$. So, $\sigma'' Ft{:}\varphi \notin S_\omega$. Thus, for any σ'', $\sigma' R\sigma''$ implies $\sigma'' Ft{:}\varphi \notin S_\omega$. Hence $\varphi \in \mathcal{E}(\sigma', t)$.

(B) We show only $\cdot$ case for closure conditions as a representative case. (Constant specification is also similar.)

Suppose $\varphi \to \psi \in \mathcal{E}(\sigma, t)$ and $\varphi \in \mathcal{E}(\sigma, t)$. Also, suppose $\sigma R\sigma'$. By definition of $\mathcal{E}$, $\sigma R\sigma' \Rightarrow \sigma' Ft{:}\varphi \to \psi \notin S_\omega$ and $\sigma R\sigma' \Rightarrow \sigma' Fs{:}\varphi \notin S_\omega$. So, $\sigma' Ft\varphi \to \psi \notin S_\omega$ and $\sigma' Fs{:}\varphi \notin S_\omega$. So, $S_\omega \cup \{\sigma' Ft{:}\varphi \to \psi\}$ and $S_\omega \cup \{\sigma' Fs{:}\psi\}$ are both S4LP-inconsistent. So, there are closed tableaux for finite subsets of these. However, out of these closed tableaux, we can construct another closed tableau having the two tableaux as branches, starting from $\sigma' Ft \cdot s{:}\psi$ using only formulas from S_ω. So, $S_\omega \cup \{\sigma' Ft \cdot s{:}\psi\}$ is S4LP-inconsistent. By S4LP-consistency of S_ω, $\sigma' Ft \cdot s{:}\psi \notin S_\omega$. ⊠

Next, we need to show the Truth Lemma. We show only relevant cases.

Lemma 4.1.36 (Truth Lemma) *For every signed formula* Φ,

$\sigma\Phi \in S_\omega \Longrightarrow \sigma$ *realizes* Φ *in the model* $\mathcal{K}$.

Proof $\Box\varphi$ cases are similar to S4LPN. The cases of $t\!:\!\varphi$ need some modification. Case (1) Suppose $\sigma Tt\!:\!\varphi \in S_\omega$. We first show the first part $\varphi \in \mathcal{E}(\sigma, t)$, which is equivalent to $\forall\sigma' \in K\ (\sigma R\sigma' \Rightarrow \sigma' Ft\!:\!\varphi \notin S_\omega)$. To show this, fix $\sigma' \in K$, $\sigma R\sigma'$, which is $\sigma \leq \sigma'$. Suppose further that $\sigma' Tt\!:\!\varphi \notin S_\omega$. Then by maximal consistency of S_ω, $S_\omega \cup \{\sigma' Tt\!:\!\varphi\}$ is inconsistent. Take a closed tableau of a finite subset S_1 of the set. (Like the case of proposition 4.1.7, w.l.o.g., we can assume that there exists a formula Φ, s.t. $\sigma'\Phi \in S_\omega$ and $\sigma'\Phi \in S_1$ s.t. $\Phi \neq Tt : \varphi$. σ' occurs in $S_1 \setminus \{\sigma' Tt : \varphi\}$, so we can use any initial segment of σ' in constructing a tableau). Since $\sigma \leq \sigma'$, by applying E4-rule repeatedly, we can produce another closed tableau for a finite subset $(S_1 \setminus \{\sigma' Tt : \varphi\}) \cup \{\sigma Tt\!:\!\varphi\}$. Since this set is included in S_ω, this is contradictory to the maximal consistency of S_ω. So, $\sigma' Tt\!:\!\varphi \in S_\omega$. By consistency, $\sigma' Ft\!:\!\varphi \notin S_\omega$. So, $\varphi \in \mathcal{E}(\sigma, t)$.

A proof of the second part, i.e. $\forall\sigma' \in K(\sigma R\sigma' \Rightarrow \sigma' \Vdash \varphi)$, is as follows. Suppose $\sigma Tt : \varphi \in S_\omega$. Assume (for contradiction) there exists $\sigma' \in K$, s.t. $\sigma \leq \sigma'$ and $\sigma' T\varphi \notin S_\omega$. (Call it σ'_1.)

We first claim $\sigma'_1 Tt : \varphi \in S_\omega$. Suppose not, i.e. $\sigma'_1 Tt : \varphi \notin S_\omega$. $S_\omega \cup \{\sigma'_1 Tt : \varphi\}$ is inconsistent. Then there exists a finite subset S_1 s.t. $S_1 \subseteq S_\omega \cup \{\sigma'_1 Tt : \varphi\}$ and S_1 has a closed tableau. (Since $\sigma'_1 \in K$, there exists some ϕ such that $\sigma'_1\Phi \in S_\omega$. As we did in the proof of proposition 4.1.7 above, let $\sigma'_1\Phi$ be in the set S_1 w. l. o. g.) By adding several steps of

E4 starting from $\sigma Tt : \varphi$ on top of the tableau, we can construct a closed tableau for $(S_1 \setminus \{\sigma'_1 Tt : \varphi\}) \cup \{\sigma Tt : \varphi\}$. However, this set is a subset of S_ω, which contradicts consistency of S_ω. Hence, $\sigma'_1 Tt : \varphi \in S_\omega$.

Now we claim $S_\omega \cup \{\sigma'_1 T\varphi\}$ is consistent. Suppose otherwise. Then there exists a finite S_1 s.t. $S_1 \subseteq S_\omega$ and S_1 has a closed tableau. By adding $\sigma'_1 Tt : \varphi$ on top of the tableau and by applying ET, we can construct a closed tableau for $(S_1 \setminus \{\sigma'_1 T\varphi\}) \cup \{\sigma'_1 Tt : \varphi\}$. However, this set is included in S_ω. Hence, it contradicts consistency of S_ω. Hence, by maximality, $\sigma'_1 T\varphi \in S_\omega$. Therefore, we have proven for any $\sigma' \in K$, $\sigma \leq \sigma'$ implies $\sigma' T\varphi \in S_\omega$. By IH, $\sigma' \in S_\omega$ implies $\sigma' \Vdash \varphi$. Thus, $\forall \sigma' \in K(\sigma R \sigma' \Rightarrow \sigma' \Vdash \varphi)$.

Case (2). Suppose $\sigma Ft : \varphi \in S_\omega$. It suffices to show that $\varphi \notin \mathcal{E}(\sigma, t)$, i.e., there exists σ', s.t. $\sigma R \sigma'$ and $\sigma' Ft{:}\varphi \in S_\omega$. But this follows from $\sigma R \sigma$ $(\sigma \leq \sigma)$ and the assumption. ⊠

Theorem 4.1.37 (Completeness) *If φ is S4LP-valid, then φ has a proof using the tableau rules for S4LP.*

Proof Similar to the case of S4LPN. ⊠

Note that the proof of completeness of the prefixed tableau system for S4LP does not use cut. Therefore, by using the standard method of proving cut-admissibility, since cut is a sound rule with respect to our semantics, the following corollary follows.

Corollary 4.1.38 *Cut is admissible in the prefixed tableau system for S4LP.*

Now we intoduce a hypersequent calculus for S4LP. First of all, a hypersequent calculus for S4LP is obtained by keeping all the rules (including cut) except modal rules and Labeled Splitting, by removing Labeled Splitting and by modifying **R**$\Box$ for S4$\Box$ into the following form.

$$\mathbf{R}\Box \; \frac{G|\Box\Gamma_1, \vec{t}{:}\Gamma_2 \Rightarrow A}{G|\Box\Gamma_1, \vec{t}{:}\Gamma_2 \Rightarrow \Box A}$$

Under the same interpretation of hypersequents as used above,

$$\mathcal{I}(\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_n \Rightarrow \Delta_n) = \Box(\bigwedge\Gamma_1 \to \bigvee\Delta_1) \vee \cdots \vee \Box(\bigwedge\Gamma_n \to \bigvee\Delta_n)$$

We can prove deductive equivalence between S4LP and HS4LP.

Theorem 4.1.39 *HS4LP* $\vdash \Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_n \Rightarrow \Delta_n$ *if and only if S4LP* $\vdash \Box(\bigwedge\Gamma_1 \to \bigvee\Delta_1) \vee \cdots \vee \Box(\bigwedge\Gamma_n \to \bigvee\Delta_n)$.

Proof In both ways, a proof is given by the length of derivation in two systems, respectively. ⊠

Now we discuss cut-admissibility of HS4LP. The way we prove cut-admissibility of HS4LP is the same as that of HS4LPN. We translate a prefixed tableau proof into a hypersequent proof. Since S4LP has no reverse rules, we do not have to specifically handle the case of reused prefixes. This implies that proof confluence is sufficient to prove that any prefixed tableau proof can be transformed into a prefixed tableau proof in a normal form, which is translatable into a hypersequent proof. Note that in S4LP, a formula with $Tt : \varphi$ allows only forward movement, so it is like a formula of the form $T\Box\varphi$. Our translation function s is similar to the case of S4LPN. We modify the notion of π-$\nu_{K,4}$ normal form (into π-$\nu_{K,4}, EK, E4$ normal form) and $\pi^\sharp$ rule to accommodate differences between S4LPN and S4LP.

Definition 4.1.40 *A prefixed tableau proof is in* π-$\nu_{K,4}, EK, E4$ *normal form if for any application of* $\nu_{K.4}$*,* $EK, E4$ *rules, there is an application of* π *rule immediately precedes it such that the application of* π *introduces a new prefix that is used in these applications. (The meaning of "immediately" is the same as before, i.e. between these* π *and* ν_{K4}*, EK, E4 only other applications of* $\nu_{K,4}$ *and EK, E4 are allowed.)*

Our $\pi^{\sharp *}$ rule for S4LP is as follows. Let S be a set of prefixed formulas, let $\sigma F\Box\varphi$ be the premise of π-rule, and let $\sigma.nF\varphi$ ($\sigma.n$ is new) be the conclusion of π-rule.

Let $S^{\sharp *}_{\sigma.n} = \{\sigma.nT\Box\psi, \sigma.nT\psi, \sigma.nTt : \rho, \sigma.nT\rho | \sigma T\Box\psi, \sigma Tt : \rho \in S\}$.

$$\frac{\{\sigma F\Box\varphi\} \cup S}{\{\sigma.nF\varphi\} \cup S^{\sharp *}_{\sigma.n}}$$

We call the system with $\pi^{\sharp *}$ TS4LP$^\sharp$. As we did in TS4LPN, we show that $\nu_{K,4}$ and $EK, E4$ rules are admissible in the system without these rules TS4LP$^{\sharp\circ}$.

Proposition 4.1.41 *If φ has a prefixed tableau proof in TS4LP, then φ has a prefixed tableau proof in TS4LP$^{\sharp\circ}$.*

Proof Similar to the case of TS4LPN$^{\sharp\circ}$. ⊠ (proposition)

We take the following steps in our translation.

1. Suppose φ has a prefixed tableau proof.

2. Pick an arbitrary prefixed tableau proof.

3. Modify the prefixed proof into a π-$\nu_{K,4}, EK, E4$ normal form (via proof confluence in TS4LP.

4. Translate the proof into an auxiliary system TS4LP$^{\sharp\circ}$.

5. Translate the proof in TS4LP$^{\sharp\circ}$ into a hypersequent calculus HS4LP.

Lemma 4.1.42 *Let S be a finite set of prefixed signed formulas. If there is a closed tableau for S using the prefixed tableau system for S4LP, then the hypersequent S^s is provable in HS4LP$^-$.*

Proof Since we have no reverse rules and EK and E4 are both absorbed in $\pi^\sharp$ rule in TS4LP$^{\sharp\circ}$, we do not have these cases. Then all the rules except $\pi^{\sharp *}$ are the same as the rules in TS4LPN. Thus, here we show only $\pi^{\sharp *}$ case.

$\sigma F \Box \varphi \in S$, and $S^{\sharp *}_{\sigma.n} = \{\sigma.nT\Box\psi, \sigma.nT\psi, \sigma.nTt : \rho, \sigma.nT\rho | \sigma T \Box \psi, \sigma Tt : \rho \in S\}$. So, $(S^{\sharp *}_{\sigma.n})^s = \Gamma^{\sharp *}_{\sigma.n}$. Then by IH, $(S \cup \{\sigma.nF\varphi\} \cup S^{\sharp *}_{\sigma.n})^s$, which is identical to $\Gamma_1 \Rightarrow \Delta_1 | \dots | \Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi | \Gamma^{\sharp *}_{\sigma.n} \Rightarrow \varphi | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$, is provable in HS4LP$^-$. We can derive S^s from the above hypersequent in HS4LP$^-$ as follows. (Here we write $\Box\psi_1, \psi_1, \dots, \Box\psi_k$ as $\overrightarrow{\Box_i\psi_i, \psi_i}$ and $t_1 : \rho_1, \rho_1, \dots, t_l : \rho_l, \rho_l$ as $\overrightarrow{t_i : \rho_i}$.)

$$
\begin{array}{cl}
\underline{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\Gamma^{\sharp*}_{\sigma.n} \Rightarrow \varphi|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}} & \text{def of } \Gamma^{\sharp*}_{\sigma.n} \\
\underline{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\overrightarrow{\Box_i\psi_i, \psi_i}, \overrightarrow{t_i : \rho_i} \Rightarrow \varphi|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}} & \text{L}\Box \\
\underline{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\overrightarrow{\Box_i\psi_i, \psi_i}, \overrightarrow{t_i : \rho_i} \Rightarrow \varphi|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}} & \text{(set)} \\
\underline{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\overrightarrow{\Box_i\psi_i}, \overrightarrow{t_i : \rho_i} \Rightarrow \varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}} & \pi^{\sharp*} \\
\underline{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\overrightarrow{\Box_i\psi_i}, \overrightarrow{t_i : \rho_i} \Rightarrow \Box\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}} & \text{IW} \\
\underline{\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}} & \text{(set)} \\
\Gamma_1 \Rightarrow \Delta_1|\ldots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\ldots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} &
\end{array}
$$

This suffices to show the following theorem.

Theorem 4.1.43 (Completeness) *If a formula φ is valid in the semantics for S4LP, then HS4LP$^-$ $\vdash\Rightarrow \varphi$.*

Proof Similar to HS4LPN$^-$.

As a corollary, we can semantically prove cut-admissibility for HS4LP.

Corollary 4.1.44 *If HS4LP $\vdash\Rightarrow \varphi$, then HS4LP$^-$ $\vdash\Rightarrow \varphi$*

Remark : 1. Originally, a prefixed tableau system for S4LP (call it S4LP$^+$) was formulated as the current system plus EFr. This system is sound with respect the same semantics.[21] In fact, by using the same definition of evidence function as used in the proof of completeness of S4LPN, i.e.

---

[21] The proof of the preservation of satisfiability for EFr is as follows: (Note that here we only use transitivity.) Suppose $\sigma.nFt:\varphi \in \theta$ and θ is satisfiable. In particular, $\exists\mathcal{K}$ and $\exists\mathcal{N}$, s.t. $\mathcal{N}(\sigma.n) \not\Vdash t:\varphi$. ($\sigma.n$ is used, so $\mathcal{N}$ is already defined for $\sigma.n$.) Then, (1) $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma.n), t)$ or (2) there exists v, s.t. $\mathcal{N}(\sigma.n)Rv$ and $v \not\Vdash \varphi$. By monotonicity of $\mathcal{E}$ and $\mathcal{N}(\sigma)R\mathcal{N}(\sigma.n)$ (since $\sigma \leq \sigma.n$), (1) implies that $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$. So, (1) implies $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$ or there exists v, s.t. $\mathcal{N}(\sigma)Rv$ and $v \not\Vdash \varphi$. On the other hand, (2) implies the following. Since $\sigma \leq \sigma.n$, $\mathcal{N}(\sigma)R\mathcal{N}(\sigma.n)$. Then, by transitivity, $\mathcal{N}(\sigma)Rv$. So, $\varphi \notin \mathcal{E}(\mathcal{N}(\sigma), t)$ or there exists v, s.t. $\mathcal{N}(\sigma)Rv$ and $v \not\Vdash \varphi$. So, either way, we get $\mathcal{N}(\sigma) \not\Vdash t:\varphi$. But this implies $Ft:\varphi$ is realized at $\mathcal{N}(\sigma)$ in $\mathcal{K}$. So, $\sigma Ft:\varphi$ is satisfiable.

$\varphi \in \mathcal{E}(\sigma, t)$ iff $\sigma F t : \varphi \notin S_\omega$, we can prove completeness of $\mathsf{TS4LP}^+$ without cut. However, if we have only E4 and EFr, then it seems impossible to find out a rule in a hypersequent calculus that exactly corresponds to these rules. So, our translation method does not work for the system.

2. Renne [135] announced that cut is admissible for a destructive tableau system for S4LP, which has a simple correspondence with an ordinary sequent calculus. It turns out that Renne's proof in [135] has an omission. Fortunately, the claim itself has turned out to be correct. Evan Goris (in personal communication) showed us that, by using the definition of evidence function that we used in our proof of completeness of S4LP, we can indeed prove completeness of the destructive tableau system for S4LP. The proof that was shown above is an adaptation of Goris' idea to the prefixed tableau system.[22] Thus, S4LP does have a cut-free destructive tableau system and an ordinary sequent calculus. So, the prefixed tableau system and the hypersequent calculus for S4LP may look redundant. From the point of view of deductive power, it is indeed the case. However, from a methodological point of view, it may have its own advantage. Unlike our older system for S4LP with EFr, our method of translation works not only for S4LPN but also S4LP. We do have certain amount of methodological uniformity in hypersequent formulations of S4LPN and S4LP.

---

[22] Partly because of fairness to Renne and partly because of its intrinsic interest, we would like to report that he also has had another correct proof of completeness of a destructive tableau system (without cut) for S4LP that uses minimal evidence function constructed out of the canonical model construction in the ordinary sense (in unpublished manuscript).

Discussion:

1. Proving completeness theorem for a hypersequent calculus for modal logic via a translation between a prefixed tableau system and a hypersequent calculus originates in Fitting ([59]). Fitting's proof is given for the case of S5, so the translation itself is straightforward. Although the translation method was not used for the purpose of proving completeness, there are some precursors of a translation like this to some different proof-theoretic framework in non-classical logics. One example is G. Mints' work ([105]) that translates indexed sequents calculi into display logics. Since our approach of using translation to prove completeness itself has not been often used, it is an interesting question how far our method can be applied to prove different kinds of non-classical logics.

2. In spite of its cut-admissibility, the hypersequent calculus HS4LPN is not analytic, i.e. does not enjoy the subformula property, due to the rule $\mathbf{R}\cdot$. So, the hypersequent calculus may not be useful for automated reasoning. We agree that the issue of the subformula property must be one of the most important open problems in the study of proof systems in justification logics. However, a complete solution to the problem is beyond the scope this section. Also, handling the issue of negative introspection in the framework is probably an issue independent of that of $\mathbf{R}\cdot$ and it is quite likely that our use of hypersequent may be reasonably combined with any approach that can handle $\mathbf{R}\cdot$. Hence, we leave this issue of the subformula

property for future research.

4.2 Tableaux and hypersequents for logic of proofs and provability

Provability logic has been used to study the notion of 'formal provability' in formal arithmetic such as PA in the concise language of modal logic. Gödel-Löb logic (GL) has been the fundamental modal logic in this area (see, [27]). On the other hand, Logic of Proofs (LP) was introduced (for instance, in [6]) as an explicit modal logic to study the structure of 'proofs' at the level of propositional logic. Yet another logic has been studied to capture the notion of "being provable in PA and true." (Such a concept of provability is called "strong provability.") The modal logic for strong provability is known as Grz [27].[23]

What kind of logic can be produced if we combine LP with these logics GL and Grz? This is a natural question. In fact, logics that combine GL (Grz) and LP have already been introduced, and their arithmetic completeness has been proven. Yavorskaya's logic LPP [178] and Artemov-Nogina's LPGL [7] are known as different versions of Hilbert-style axiomatic systems combining GL and LP. (We go back to differences of these systems later.) Also, a system that combines Grz and LP is given in Nogina [109]. Other complete

---

[23] This is sometimes referred to as S4Grz, but all the axioms except what is commonly called K axiom can be derived by the characteristic axiom for Grz. So, let us call the logic Grz here.

axomatizations of these logics with respect to Fitting-style semantics are given under the name GLA and GrzA ([109], [110]). The general interest of combining these logics consists in the following. GL itself can be an interesting tool to consider what kind of sentences containing a provability predicate in PA by using the concise language of modal logic. By combining GL and LP, we can extend this observation to the interaction between the notion of formal provability in PA and the notion of proofs in PA. A good illustration of this point is actually one of the axioms connecting a formula with a proof-term and modality, which has the form $\neg t : \varphi \to \Box\neg t : \varphi$. This statement is not only intuitively correct under an epistemic interpretation but it is also valid in arithmetic interpretation. (For more concrete examples of arithmetically valid formulas on which we can play around by using this combined language, see Yavorskaya's [178].) This formula, which occurs as an axiom of S4LPN, a logic combining LP and S4 via some connecting axioms (one of them is the mixed negative introspection), is also of interest from the perspective of (epistemic interpretation of) modal logic, since it can be taken as a kind of negative introspection (we call this "mixed negative introspection"), which is an analogue of 5 axiom $\neg\Box\varphi \to \Box\neg\Box\varphi$ in S5 modal logic, and $\neg t : \varphi \to \Box\neg t : \varphi$ can even be better motivated than the negative introspection $\neg\Box\varphi \to \Box\neg\Box\varphi$. This is due to a broader consideration about the nature of evidence. The underlying idea of LP [6] enables us to take a positive occurrence of $\Box\varphi$ to mean something like an implicit existential statement quantifying over some "evidence" for φ.[24] From this point of view, $\neg\Box\varphi$ is a potentially universal statement which states that there is no evidence for φ. Then, to establish the truth of the statement $\Box\neg\Box\varphi$, for (possibly) infinitely many candidates of evidence of φ, we may have to give an evidence that these are not evidences for φ. This does not sound very plausible. On the other hand, we need only to verify that a particular evidence given by a proof-term t does not work for φ for the mixed introspection case.

In this section, we introduce prefixed tableau systems and hypersequent calculi for these two logics GLA and GrzA. The main goal of this section is to prove cut-admissibility of hypersequent calculi HGLA and HGrzA via a semantic method.

But why do we need to use hypersequents for GLA and GrzA? The main difficulty to formulate a cut-free traditional sequent calculus for GLA (GrzA) is due to the very formula that we have just discussed, namely $\neg t : \varphi \to \Box\neg t : \varphi$ ($\neg t : \varphi \to \boxdot\neg t : \varphi$). It is well known that formulating an ordinary sequent calculus (or a destructive tableau system) for S5 is difficult, and the mixed negative introspection axiom for S5 shares its difficulty in proof-theory with S5. In the case of S5, there are at least two commonly known (and several less commonly known) methods of formulating proof systems for the logic. One is to use some label representing a possible world in a proof

---

[24] This can be taken as an informal "existential semantics" for modality in contrast to traditional Kripke semantics, which is like a "universal semantics," since the underlying idea of interpreting $\Box$ modality uses universal quantifier, i.e., truth in every accessible world.

system. Above all, Fitting's prefixed tableau system [57], which introduces finite sequences of positive integers (prefixed to formulas) representing possible worlds, can formulate many modal logics including S5, and the other is to use a hypersequent calculus to formulate S5. Avron [18] formulated a cut-free hypersequent calculus for S5. (For comparisons among different proof systems in modal logics, see Chapter 2.) Thus, it is a natural conjecture that a similar idea may work for this mixed negative introspection axiom.

Although there is an obvious similarity among GLA, GrzA, S5, and S4LPN, there are some differences which we now want to highlight. Let us start discussing a difference between S5 and all other logics containing LP. A semantic proof of cut-admissibility for a hypersequent calculus for S5 is not too complicated. We can construct a saturated hypersequent only by using subformulas of an unprovable hypersequent and turn this into a countermodel of it. The idea is given in Avron [18]. In this sense, in order to prove cut-admissibility, one may *not have to* use prefixed tableau system and a hypersequent calculus, although Fitting [59] presented a connection between these two systems and use the connection to prove completeness for the cut-free hypersequent calculus. However, the situation in S4LPN, GLA and GrzA is not so simple. In any logic having proof-terms in LP style, we need to handle closure conditions for an evidence function like the case of S4LPN, so directly proving completeness for a cut-free hypersequent calculus would be very complicated. Hence, connecting these two approaches (namely prefixed tableaux and hypersequents) applying the method of a translation given by Fitting [59] to the case of S5 may significantly facilitate a proof of completeness. So, we are applying essentially same technique to GrzA and GLA.

However, GLA and GrzA have two further problems. Since the case of GrzA is simpler, we first explain the difference between S4LPN and GrzA. The frame condition for GrzA contain the condition that there is no infinite strictly ascending chain. To satisfy this condition, as is well-known, the ordinary canonical model construction does not work. We have to "finitize" the construction to satisfy the frame condition. In the case of Grz itself, one can avoid taking infinitary construction of maximal consistent set since it suffices to take maximal consistent from the subformulas of the given formula whose countermodel we are suppose to construct. However, for GrzA, we also need to construct an evidence function that satisfies the closure conditions. Hence, it is very unlikely that we can simply avoid infinitary construction (although there may be slightly more economical method than ours). Here we use a two-stage method for proving completeness of GrzA (and GLA). First, we construct a canonical model for GrzA by the so-called Lindenbaum-Henkin construction presented in [59]. Then, by using the notion of n-bisimulation (a bisimulation bounded by the complexity of a formula whose countermodel we try to construct), we construct another model whose height is at most n where n is a modal depth of the pertinent formula and that is n-bisimilar to the canonical model for the prefixed tableau system for GrzA. Since the canonical model construction is sufficient to provide an evidence function, forcing relation, etc., the frame condition that there is no strictly ascending chain is now clearly satisfied for the smaller model, and the two models satisfy the same formula up to the depth n, the smaller model works as a countermodel for the original formula. This essentially suffices to prove cut-free completeness for the prefixed tableau system for GrzA. Once this is done, we can apply the aforementioned translation method to a hypersequent calculus for GrzA to prove cut-free completeness for the hypersequent calculus for GrzA.

Lastly, in order to show completeness of a cut-free prefixed tableau system for GLA, we have to overcome the remaining problem. A complication of GLA essentially comes from its special rule called "reflection rule," which is as follows.

$$\frac{\Box\varphi}{\varphi}.$$

Note that an axiom schema $\Box\varphi \to \varphi$ is, of course, inconsistent with GL, but the rule version does not produce any inconsistency. Moreover, this rule turns out to be useful if we add this rule to a variant of GL.[25] For some technical reason [7], it is desirable to have this rule in our system, too, although adding this obviously prevents the system from having the subformula property, but not in a very harmful manner.[26]

---

[25] The rule is used in [4], but the rule was introduced in the literature of provability logic by Guraspari and Solovay when they formulated a provability logic for Rosser provability in [78].

[26] We will discuss the issue later, but all currently available Gentzen-style systems for logics in LP family (justification logics) have a more serious problem of the application

To accommodate this rule in a prefixed tableau system, we first formulate a prefixed tableau system for a subsystem for GLA without reflection rule (we call the subsystem wGLA).[27] We prove soundness and completeness of the prefixed tableau system for wGLA. Adding the new rule, we prove the soundness of the prefixed tableau system for GLA. For completeness for the prefixed tableau system for GLA, we consider a model of wGLA and make modifications of the model so that the model validates the Reflection Rule, i.e. becomes a model of GLA. This involves a semantic construction originally due to Guaspari and Solovay [78] (our direct source is [4]). Finally, we consider a translation from the new prefixed tableau system for GLA to a new hypersequent calculus.

4.2.1 Hilbert-style axiomatic systems and Kripke models for **GLA** and **GrzA**

The language of GLA can be specified as follows:

1. The class of proof terms in the language of GLA is specified as follows.

---

rule in LP. Compared with this problem, this violation is a minor one. A further reason to think that this is relatively benign will be explained later.

[27] To be precise, we also omit one axiom $t : \Box\varphi \to \varphi$ in order to talk about the subsystem. However, the semantic condition that validate the axiom is actually the same as that of validating reflection rule, so we discussed only reflection rule in the main text. Again, adding this axiom has a technical advantage. We will discuss the issue in the discussion section towards the end of this paper when we compare the system in [7] and the system in [178].

$t := x|a|!t|t_1 \cdot t_2|t_1 + t_2.$

2. The class of formulas in the language of GLA is specified as follows:

$$A := p_i|\bot|\neg A_1|A_1 \rightarrow A_2|A_1 \wedge A_2|A_1 \vee A_2|t : A|\Box A$$

Note: In proof-terms, x is a proof variable, a is a proof constant, and we have proof operators: !(proof-checker) , $\cdot$ (proof application), + (proof sum). (For arithmetical meaning of these operations, see [6].)

In this section, “Trm” stands for the set of proof-terms in the language of GLA or GrzA. “Fmla” stands for the set of formulas in the language of S4LPN or GrzA. (The language of GrzA will be officially defined shortly. Unless we have a possibility of confusion, we do not specify which we mean.)

A *constant specification* is a mapping $\mathcal{CS}$ from proof constants to sets of formulas (possibly empty). A formula A *has a proof constant* c with respect $\mathcal{CS}$ if $A \in \mathcal{CS}(c)$.

Hilbert style system of GLA is as follows.

Axioms:

0) Axioms of Propositional Logic

1) Axioms of Logic of Proofs:

1. $t : F \to F$ 2. $t : (F \to G) \to (s : F \to t \cdot s : G)$

3. $t : F \to !t : t : F$ 4. $t : F \to t + s : F$, $s : F \to t + s : F$

2) Axioms of Logic of Provability:

1. $\Box(F \to G) \to (\Box F \to \Box G)$ 2. $\Box F \to \Box\Box F$

3. $\Box(\Box F \to F) \to \Box F$

3) Connecting Axioms:

1. $t : F \to \Box F$ 2.$\neg t : F \to \Box\neg t : F$ 3. $t : \Box F \to F$

Rules of Inference:

1. Modus Ponens 2. Necessitation $\dfrac{F}{\Box F}$ 3. Reflection Rule $\dfrac{\Box F}{F}$

4. $c : A$ (Axiom Necessitation), where A is one of the above axioms of GLA and $A \in \mathcal{CS}(c)$

Note: For some technical reason, we define an auxiliary subsystem of GLA that we can obtain by removing the connecting axiom 3. $t : \Box F \to F$ and reflection rule (and axiom necessitation is weakened accordingly). We call the system wGLA.

We now define Fitting-style Kripke semantics for wGLA. Before moving on, let us mention our notational convention. We sometimes use $\Longrightarrow$ as an abbreviation of "if . . . , then . . . " in the metalanguage, which is intended to be a classical implication in the metalanguage. Also, occasionally we use "$\forall$" and "$\exists$" as abbreviations of English phrase "for all" and "there exists". Since we do not use quantifiers in our object language, there should not be any possibility of confusion.

Let a quadruple (K, R, R^e, r) be a frame, where K is non-empty set. It is, in general, sufficient to specify R as an irreflexive and transitive relation on K that has no infinite ascending chain. However, here we take R to be a strict partial order with a single root r that has no infinite ascending chain. Note that this implies that $\forall u \in K, r \neq u \Longrightarrow rRu$. R^e is a reflexive, symmetric and transitive relation. We have an additional condition $R \subseteq R^e$.

Let $\mathcal{E}$ be an evidence function: $K \times Trm \longrightarrow \mathcal{P}(Fmla)$ that satisfies the following properties (A constant specification CS is a function from proof constants to the set of axioms in wGLA):

1. uR^ev implies $\mathcal{E}(u,t) \subseteq \mathcal{E}(v,t)$ (Monotonicity)

2. $F \to G \in \mathcal{E}(u,t)$ and $F \in \mathcal{E}(u,s)$ implies $G \in \mathcal{E}(u, t \cdot s)$
3. $F \in \mathcal{E}(u,t)$ implies $t : F \in \mathcal{E}(u, !t)$
4. $\mathcal{E}(u,s) \cup \mathcal{E}(u,t) \subseteq \mathcal{E}(u, s+t)$
5. $\mathcal{CS}(c) \subseteq \mathcal{E}(u,c)$

Note: In fact, $uR^e v$ implies $\mathcal{E}(u,t) = \mathcal{E}(v,t)$ holds. This is due to symmetry of R^e. We call this *stability*.

Then, we define a Kripke GLA model $\mathcal{K}$, which is actually a model only for wGLA, but we call it a GLA model. (We explain the reason right after giving the definition.) A GLA model is defined as a sextuple $(K, R, R^e, r, \mathcal{E}, \Vdash)$, where r is the root node of the relation R and $\Vdash$ is a forcing relation that satisfies the following conditions:

1. $\Vdash$ commutes with Booleans at each state; for all $u \in K$, $u \nVdash \bot$.
2. $u \Vdash \Box\varphi$ iff for every $v \in K$, s.t. uRv, $v \Vdash \varphi$
3. $u \Vdash t : \varphi$ iff $\varphi \in \mathcal{E}(u,t)$ and for every $v \in K$, s.t. $uR^e v$, $v \Vdash \varphi$.
4. $A \in \mathcal{CS}(c)$ implies $\mathcal{K}, u \Vdash c : A$ for every $u \in K$.

These are the conditions of a GLA model (a model of wGLA). However, we have one additional condition for GLA itself.

Root Soundness : Let r be the root of a tree GLA model. Then $r \Vdash Rf(Sb(\varphi))$ for a given formula φ, where $Rf(Sb(\varphi)) = \{\Box A \to A | \Box A \in$

$Sb(\varphi)\}$. (Rf stands for "root formulas".) We call a model that satisfies this condition "a φ-sound GLA model". In addition, since we always have a constant specification $\mathcal{CS}$ for GLA, officially we consider φ-sound GLA model with $\mathcal{CS}$ for GLA. (We invariably use the full constant specification for GLA in this paper.)

Note: Although we call the defined model (without using the Root Soundness condition) GLA model, only wGLA axioms and rules are sound with respect to the class of GLA models defined in this way. The axiom $t : \Box\varphi \to \varphi$ and reflection rule $\dfrac{\Box\varphi}{\varphi}$ are sound only with respect to $\Box\varphi$-sound GLA models. Unfortunately, this terminology seems to be inevitable, since $\Box\varphi$-soundness depends on a particular formula $\Box\varphi$ whose models we want to consider. We cannot give a general definition of a model for GLA. The same comment applies to completeness, but the latter is not an unbearably severe restriction since GL is not strongly complete anyway so that we have to state the statement of weak completeness only for a single formula.

When for any $u \in K$, $u \Vdash \varphi$, φ is said to be *valid in a model* $\mathcal{K}$, and we use the notation $\mathcal{K} \Vdash \varphi$.

Now we present a Hilbert-style axiom system for GrzA and Kripke semantics.

The language of GrzA can be specified as follows:

1. The class of proof terms in the language of GrzA is specified as follows.

$t := x|a|!t|t_1 \cdot t_2|t_1 + t_2$.

2. The class of formulas in the language of GrzA is specified as follows:

$A := p_i|\bot|\neg A_1|A_1 \rightarrow A_2|A_1 \wedge A_2|A_1 \vee A_2|t : A| \boxdot A$

A Hilbert style system of GrzA is as follows.

Axioms: 0) Axioms of Propositional Logic

1) Axioms of Logic of Proofs: same as GLA

2) Axioms of Logic of Strong Provability: 1. $\boxdot(F \rightarrow G) \rightarrow (\boxdot F \rightarrow \boxdot G)$

2. $\boxdot F \rightarrow F$ 3. $\boxdot F \rightarrow \boxdot \boxdot F$ 4. $\boxdot(\boxdot(F \rightarrow \boxdot F) \rightarrow F) \rightarrow F$

3) Connecting Axioms: 1. $t : F \rightarrow \boxdot F$ 2. $\neg t : F \rightarrow \boxdot \neg t : F$

Rules of Inference: 1. Modus Ponens 2. Necessitation $\dfrac{F}{\boxdot F}$

3. $c : A$ (Axiom Necessitation), where A is one of the axioms of GrzA and

$A \in \mathcal{CS}(c)$.

A Kripke model $\mathcal{K}^s(= (K^s, R^s, R^e, r, \mathcal{E}, \Vdash))$ for GrzA can be described as follows.

1. R^s is a non-strict partial order (with a single root) that has no strictly ascending infinite chain.[28]

2. Other conditions of Kripke models for GLA and GrzA are the same.

Note that the root soundness condition is redundant for GrzA, since a GrzA-model is reflexive. As a result, we do not have to specify the root node r here for a GrzA model. Nonetheless, our prefixed tableau system is naturally connected to a single-rooted model, so w.l.o.g., here we consider such a model for GrzA.

4.2.2 Prefixed Tableau System for **GrzA** and **GLA**

For the aforementioned expository reason, we first introduce a prefixed tableau system for GrzA (TGrzA) and we prove soundness and completeness for TGrzA.

4.2.2.1 Prefixed Tableau System for **GrzA**

Here is the prefixed tableau system for GrzA. We call it TGrzA. In the following, we follow Fitting's terminology for basic notions in the prefixed tableau

---

[28] See, [32]. The last condition is equivalent to weakly converse well-foundedness by using ZF plus Axiom of Dependent Choice [89].

system and we omit some details, taking them for granted (e.g., branch, tableau, closed branch, closed tableau). We call a formula $1F\varphi$ having a tableau proof in TGrzA if it has a closed tableau in TGrzA. Our prefix is a finite sequence of natural numbers that has only 1 for the initial element of a sequence of natural numbers. In $\sigma T\varphi$ or $\sigma F\varphi$, σ is a prefix of a signed formula $T\varphi$ or $F\varphi$, where T and F are intended to mean "true" and "false" respectively, and a prefix σ is a sequence of positive integers, intended to represent a possible world in a Kripke model. We say σ is *used* on a branch if a prefix that has σ as its (not necessarily proper) initial segment has already occurred on the branch, and we say σ is new, otherwise. (We call these "preconditions" of prefixes for applications of rules.) $\sigma.n$ is called a simple extension of σ. σ' is accessible from σ iff $\sigma \leq \sigma'$ ($\leq$ means that "is an (not necessarily proper) initial segment of"). σ' is e-accessible from σ iff $1 \leq \sigma$ and $1 \leq \sigma'$. (I.e., for any non-empty sequence σ, σ', σ is e-accessible to σ'. Note that such e-accessibility is an equivalence relation.)

Rules for propositional logic

Rules for LP: (ν-rules for explicit proofs)[29]

$$\text{EK}\ \frac{\sigma Tt:\varphi}{\sigma.nT\varphi}\ (\sigma.n \text{ is used.}) \qquad \text{ET}\ \frac{\sigma Tt:\varphi}{\sigma T\varphi}$$

---

[29] In E4r and EFr, we take any subsequence of $\sigma.n$ occurs in it to have already occurred in the pertinent branch of a tableau. So, we do not need any restriction.

E4 $\dfrac{\sigma Tt:\varphi}{\sigma.nTt:\varphi}$ ($\sigma.n$ is used.)

E4r $\dfrac{\sigma.nTt:\varphi}{\sigma Tt:\varphi}$ EF $\dfrac{\sigma Ft:\varphi}{\sigma.nFt:\varphi}$ ($\sigma.n$ is used.) EFr $\dfrac{\sigma.nFt:\varphi}{\sigma Ft:\varphi}$

Operational Rules on F's : $\cdot$-rule $\dfrac{\sigma F(t\cdot s):\varphi}{\sigma Ft:\psi\to\varphi|\sigma Fs:\psi}$

!-rule $\dfrac{\sigma F!t:t:\varphi}{\sigma Ft:\varphi}$ +-rule $\dfrac{\sigma F(s+t):\varphi}{\sigma Ft:\varphi}$ $\dfrac{\sigma F(t+s):\varphi}{\sigma Ft:\varphi}$

Modal Rules:

ν-rules: ν_K $\dfrac{\sigma T\boxdot\varphi}{\sigma.nT\varphi}$ ($\sigma.n$ is used.) ν_4 $\dfrac{\sigma T\boxdot\varphi}{\sigma.nT\boxdot\varphi}$ ($\sigma.n$ is used.)

ν_T $\dfrac{\sigma T\boxdot\varphi}{\sigma T\varphi}$

π-rule for Grz: $\dfrac{\sigma F\boxdot\varphi}{\begin{array}{l}\sigma.nF\varphi\\ \sigma.nT\boxdot(\varphi\to\boxdot\varphi)\end{array}}$ ($\sigma.n$ is new.)

In addition, we have Constant Specification Rules as follows: a branch is closed if it has $\sigma Fc:A$ ($A\in\mathcal{CS}(c)$).

We give a few definitions. A signed formula $F\varphi$, $T\varphi$ (written Φ schematically) is realized at a possible world u of a model $\mathcal{K}^s$ if 1) the formula is $T\varphi$ and $\mathcal{K}^s, u \Vdash \varphi$ or 2) the formula is $F\varphi$ and $\mathcal{K}^s, u \nVdash \varphi$. To introduce the notion of a prefixed, signed formula being satisfiable, we need such a thing called "interpretation function." A (partial) function $\mathcal{N}$ from the set of σ-sequences to the set of worlds K^s is a GrzA interpretation (partial) function if, for any $\sigma, \sigma' \in Dom(\mathcal{N})$,

(1) $\exists u\ (\mathcal{N}(\sigma) R^s u)$ and σ' is accessible from $\sigma \Longrightarrow \mathcal{N}(\sigma) R^s \mathcal{N}(\sigma')$ and

(2) $\exists u\ (\mathcal{N}(\sigma) R^e u)$ and σ' is e-accessible from $\sigma \Longrightarrow \mathcal{N}(\sigma) R^e \mathcal{N}(\sigma')$.[30]

A set S of prefixed, signed formula is *satisfiable* if there is a model $\mathcal{K}^s$ and a mapping $\mathcal{N}$ from the prefixes in S to possible worlds in $\mathcal{K}^s$, such that if $\sigma\Phi \in S$, then Φ is realized at $\mathcal{N}(\sigma)$ in $\mathcal{K}^s$, where Φ is a signed formula. A tableau branch is satisfiable if the set of prefixed formulas on it is satisfiable. A tableau is satisfiable if some tableau branch is. Now we prove soundness of the prefixed tableau system.

Lemma 4.2.1 *Suppose $\mathcal{T}$ is a satisfiable tableau. If any tableau rule for GrzA is applied to $\mathcal{T}$, then the resulting tableau is still satisfiable.*

Proof Suppose a tableau is GrzA-satisfiable because a branch θ of $\mathcal{T}$ is GrzA-satisfiable, i.e. its members are realized at $\mathcal{N}(\sigma)$ of model $\mathcal{K}^s$. Suppose that a

---

[30] We consider only $\mathcal{N}$ s.t. $\mathcal{N}(1) = r$.

tableau rule for GrzA is applied to the tableau $\mathcal{T}$. The entire proof is divided into two cases:

Case 1: Our tableau rule is not applied on the branch θ.

Then, θ is still present on the new tableau and θ is satisfiable, which makes θ obviously GrzA-satisfiable.

Case 2: Since α and β cases are straightforward, we focus only the cases of ν and π. In GrzA, we have LP-rules and $\boxdot$-rules, but most cases overlap with rules in S4LPN. Thus, we treat only π-rule for Grz, whose proof is substantially complicated.

π-rule (of Grz): Suppose $\sigma F \boxdot \varphi \in \theta$ and θ is satisfiable, but $\theta \cup \{\sigma.nF\varphi, \sigma.nT \boxdot (\varphi \to \boxdot\varphi)\}$ is not satisfiable. By the definition of satisfiability, the former implies (1) $\exists\mathcal{K}^s\exists\mathcal{N}$, s.t. $\mathcal{K}^s$ is based on a frame in which R^s has no strictly ascending infinite chain, and for any $\sigma_i T\psi_i \in \theta$, $\mathcal{N}(\sigma_i) \Vdash \psi_i$ and for any $\sigma_i F\psi_i \in \theta$, $\mathcal{N}(\sigma_i) \nVdash \psi_i$, and, in particular, one of such ψ_i is $\boxdot\varphi$ and $\mathcal{K}^s, \mathcal{N}(\sigma) \nVdash \boxdot\varphi$, Also, the latter implies (2) $\forall\mathcal{K}^s, \forall\mathcal{N}$, if for any $\sigma_i T\psi_i \in \theta$, $\mathcal{N}(\sigma_i) \Vdash \psi_i$ and for any $\sigma_i F\psi_i \in \theta$, $\mathcal{N}(\sigma_i) \nVdash \psi_i$ and $\mathcal{N}(\sigma.n) \Vdash \boxdot(\varphi \to \boxdot\varphi)$, then $\mathcal{N}(\sigma.n) \Vdash \varphi$. We call $\mathcal{K}^s_1$ and $\mathcal{N}_1$ a model and an interpretation stated to exist in (1), respectively. Then, from (1) it follows that $\mathcal{K}^s_1, \mathcal{N}_1(\sigma) \nVdash \boxdot\varphi$. So, $\exists v \in \mathcal{K}^s$ ($\mathcal{N}_1(\sigma)R^s v$ and $v \nVdash \varphi$). We temporarily call the world v' Since this $\mathcal{N}_1$ is not defined for $\sigma.n$, we extend $\mathcal{N}_1$ so that $\mathcal{N}_1(\sigma.n) = v'$.[31] Suppressing

---

[31] Note that we have to check that the extended $\mathcal{N}_1$ is still an GrzA interpretation. [57] gives details. But anyways we show an argument here. Let $\mathcal{N}_1(\sigma.n) = u$. (u is a state whose existence is stated by the existential statement in the definition.) Obviously, $\sigma < \sigma.n$. This is supposed to imply $\mathcal{N}_1(\sigma)R^s\mathcal{N}_1(\sigma.n)$. This R^s has to satisfy the condition of accessibility for GrzA. 1. Reflexivity. Since uR^su by reflexivity, $\mathcal{N}_1(\sigma.n)R^e\mathcal{N}_1(\sigma.n)$.

$\mathcal{K}_1$, by (2) we get that $\mathcal{N}_1(\sigma.n) \Vdash \boxdot(\varphi \to \boxdot\varphi)$ implies $\mathcal{N}_1(\sigma.n) \Vdash \varphi$. Since $\mathcal{N}_1(\sigma.n) = v'$, $v' \Vdash \boxdot(\varphi \to \boxdot\varphi)$ implies $v' \Vdash \varphi$. Combined with $v' \nVdash \varphi$, we get $v' \nVdash \boxdot(\varphi \to \boxdot\varphi)$. So, $\exists u_1$ ($v'R^s u_1$ and $u_1 \nVdash \varphi \to \boxdot\varphi$). So, $\exists u_1$ ($v'R^s u_1$ and $u_1 \Vdash \varphi$ and $u_1 \nVdash \boxdot\varphi$) So, $\exists u_1 \exists u_2 (v'R^s u_1$ and $u_1 \Vdash \varphi$ and $u_1 \nVdash \boxdot\varphi$ and $u_1 R^s u_2$ and $u_2 \nVdash \varphi)$. [We claim $v' \neq u_1$ and $u_1 \neq u_2$ and $v' \neq u_2$. The first two are obvious, since $v' \nVdash \varphi$, $u_2 \nVdash \varphi$ but $u_1 \Vdash \varphi$. The last follows from anti-symmetry of R^s. Indeed, suppose $v' = u_2$. Then $v'R^s u_1$ and $u_1 R^s v'$. Hence, $v' = u_1$. But this is already shown to be impossible.] Let us put $\mathcal{N}_1(\sigma) = t$ from now on. By transitivity, we have $tR^s u_2$. (By the same reasoning, clearly we have $t \neq u_2$.)

We now consider another instance of $\mathcal{N}$ (call it $\mathcal{N}_2$) such that $\mathcal{N}_2(\sigma_i) = \mathcal{N}_1(\sigma_i)$ for any σ_i occurring in θ and $\mathcal{N}_2(\sigma.n) = u_2$ ($u_2 \neq u_1$) on the basis of the fact that we have derived $t \nVdash \boxdot\varphi$ and the existence of u_2 s.t. $tR^s u_2$ and $u_1 \neq u_2$. (Due to the feature that $\mathcal{N}_2$ coincide with $\mathcal{N}_1$ for prefixes occurring in θ, $\mathcal{N}_2$ can be used to show that there is another world in the model $\mathcal{K}^s$ by

---

Hence, this obviously holds. 2. No strictly ascending infinite chain. $\sigma.n$ is just one-step extension and has nothing to do with strictly ascending infinite chain. 3. Transitivity. Let $\sigma' \leq \sigma$, $\mathcal{N}_1$ is defined for σ' and $\mathcal{N}_1(\sigma')$. Hence, $\mathcal{N}_1(\sigma')R^s\mathcal{N}_1(\sigma)$. On the other hand, $\sigma' \leq \sigma$ and $\sigma \leq \sigma.n$. So, $\sigma' \leq \sigma.n$. This requires $\mathcal{N}_1(\sigma')R^s\mathcal{N}_1(\sigma.n)$. Indeed, we can prove this as follows. Since $\mathcal{N}_1(\sigma)R^s u$ and $\mathcal{N}_1(\sigma.n) = u$, $\mathcal{N}_1(\sigma)R^s\mathcal{N}_1(\sigma.n)$. Hence, by transitivity of R^s, we can obtain $\mathcal{N}_1(\sigma')R^s\mathcal{N}_1(\sigma.n)$.

R^e has to satisfy the following conditions. 1. Reflexivity. Similar to the above case. 2. Symmetry. (Clearly, $\sigma.n$ is e-accessible from σ implies σ is e-accessible from $\sigma.n$.) Suppose $\mathcal{N}_1(\sigma)R^e\mathcal{N}_1(\sigma.n)$. We show $\mathcal{N}_1(\sigma.n)R^e\mathcal{N}_1(\sigma)$. However, $\mathcal{N}_1(\sigma.n) = u$. Hence, $\mathcal{N}_1(\sigma)R^e u$. Due to symmetry of R^e itself, we get $uR^e\mathcal{N}_1(\sigma)$. Thus, $\mathcal{N}_1(\sigma.n)R^e\mathcal{N}_1(\sigma)$. 3. Transitivity: σ is e-accessible from σ' and $\sigma.n$ is e-accessible from σ. $1 \leq \sigma$ and $1 \leq \sigma'$ and $1 \leq \sigma.n$ and $1 \leq \sigma$. Hence, $\sigma.n$ is e-accessible from σ'. Suppose $\mathcal{N}_1(\sigma')R^e\mathcal{N}_1(\sigma)$ and $\mathcal{N}_1(\sigma)R^e\mathcal{N}_1(\sigma.n)$. $\mathcal{N}_1(\sigma)R^e\mathcal{N}_1(\sigma.n)$ implies $\mathcal{N}_1(\sigma)R^e u$. By transitivity, $\mathcal{N}_1(\sigma')R^e u$. But again this implies $\mathcal{N}_1(\sigma')R^e\mathcal{N}_1(\sigma.n)$.

taking the same step as we showed above and showing the existence of two new worlds in the model. Note that $\mathcal{N}_2$ does not have v' and u_1 in its range, but an interpretation function does not have to be an onto function.) Since $\mathcal{N}_2$ satisfies all the formulas from θ, we proceed as follows.

We have $\mathcal{N}_2(\sigma)R^s\mathcal{N}_2(\sigma.n)$ and $\mathcal{N}_2(\sigma.n) \nVdash \varphi$. We also take an instantiation of (2), i.e. $\mathcal{N}_2(\sigma.n) \Vdash \boxdot(\varphi \rightarrow \boxdot\varphi)$ implies $\mathcal{N}_2(\sigma.n) \Vdash \varphi$. However, since $\mathcal{N}_2(\sigma.n) = u_2$, $u_2 \nVdash \boxdot(\varphi \rightarrow \boxdot\varphi)$. So, $\exists u_3(u_2R^su_3$ and $u_3 \Vdash \varphi$ and $u_3 \nVdash \boxdot\varphi)$. (As before, we can take u_3 s.t. $u_2 \neq u_3$, $v' \neq u_3$ and $u_1 \neq u_3$.) So, further, $\exists u_4$ $(u_3R^su_4$ and $u_4 \nVdash \varphi)$, etc. (where $v' \neq u_4$ and for any $i \leq 3$, $u_i \neq u_4$). This argument can be repeated indefinitely. More precisely, we claim the following.

Claim 4.2.2 *If 1)* $u_1R^su_2 \ldots u_{n-1}R^su_n$, *2)* $\bigwedge_{1\leq i\neq j\leq n} u_i \neq u_j$ *and 3)* $u_nR^su_{n+1}$ *and* $u_n \neq u_{n+1}$, *then* $u_{n+1} \neq u_i$ *for any* i, *s.t.* $1 \leq i \leq n$ *(for any* $n \geq 2$*).*

Proof Suppose 1), 2), 3) are all true but the consequent is false, i.e., $u_{n+1} = u_i$ for some i, s.t. $1 \leq i \leq n$. We have $u_1R^su_2 \ldots u_iR^su_{i+1} \ldots u_{n-1}R^su_n$. By transitivity, $u_{i+1}R^su_{n+1}$. Also, since $u_{n+1}R^su_{n+1}$, by substituting $u_i = u_{n+1}$, $u_{n+1}Ru_i$. Since $u_iR^su_{i+1}$, by transitivity $u_{n+1}R^su_{i+1}$. By anti-symmetry, we have $u_{i+1} = u_{n+1}$. This further implies $u_i = u_{i+1}$. But this clearly contradicts 2) [and 3) when $i = n$]. So, $u_{n+1} \neq u_i$ for any i, s.t. $1 \leq i \leq n-1$. ⊠(claim)

We will then produce $v'R^su_1R^su_2R^su_3R^su_4 \ldots$ where every world in this chain is pairwise distinct. But this contradicts to the fact that there is no (non-reflexive) ascending infinite chain of R^s in any model of GrzA. ☒

Now we can state and prove the soundness theorem.

Theorem 4.2.3 (Soundness) *If φ has a prefixed GrzA-tableau proof, then φ is valid in all models for GrzA.*

Proof Suppose φ has an GrzA-tableau proof, but is not GrzA-valid. Say, φ does not hold at a world s of some GrzA-model. W.l.o.g., we can take this s to be the root state r, since we can take a generated submodel if s is not the root. (Note that $\mathcal{E}$ is stable, so a generated submodel of a model of GrzA is still a model of GrzA which preserve forcing relation of the original model at s.) Now a prefixed tableau begins with $1F\varphi$. Define an GrzA-interpretation $\mathcal{N}$ by setting $\mathcal{N}(1) = r$.

Since φ is not forced at r, i.e. $r \nVdash \varphi$, the starting GrzA-tableau is GrzA-satisfiable ($\mathcal{N}(1) \nVdash (\varphi)$), so $\{1F\varphi\}$ is GrzA-satisfiable. By the above lemma, so is every subsequent tableau. But an GrzA-satisfiable tableau cannot be closed, which contradicts the assumption that φ had a tableau proof.

Therefore, if φ has a prefixed GrzA-tableau proof, then φ must be GrzA-valid. ☒

Now we move on to the completeness theorem. We construct a countermodel for any formula unprovable in GrzA. Since the structure of the proof is quite complicated, let us give an outline of a construction of a countermodel for the unprovable formula φ in GrzA.

Step 1. We carry out Lindenbaum-Henkin construction and construct a canonical model for GrzA.

Step 2. We verify that the constructed canonical model satisfies the conditions of evidence function and forcing relation (Truth Lemma). However, our canonical model does not satisfy the property of the underlying frame of a GrzA model, i.e. there is no infinite strictly ascending chain. (In this sense, calling this canonical "model" *pseudo-model* may be appropriate.[32]) In order to construct a countermodel that satisfies the desired property, we are taking further steps.

Step 3. We define the notion of bounded bisimulation and prove a general proposition that for any formula of degree n, n-bisimilar state of the two models are equivalent.

Step 4. We construct a finite height (rooted) model from the pseudo-model $\mathcal{K}^s$ whose height is the same as the number of the degree of the formula φ whose countermodel we are constructing, and we show that the root node of $\mathcal{K}^s$ and the root node of the constructed model force the same formulas up to degree n.

Step 5. We conclude that the constructed n-height model is indeed a countermodel for φ.

Step 1: We start Lindenbaum-Henkin construction ([59]). The argument

---

[32] The word is used in [121] in the first proof of completeness of propositional dynamic logic (PDL). Although the context is slightly different, there is a similarity. We need some additional care of the frame condition of the model to be constructed.

is similar to that of [98]. We first give some definitions.

A set S of prefixed formulas is GrzA-consistent if no GrzA-tableau for a finite part of S is closed. S is maximally GrzA-consistent if S is GrzA-consistent and no S', s.t. $S \subsetneq S'$ is GrzA-consistent. S is π-complete (or $F \boxdot \varphi$-complete) provided, if $\sigma\pi \in S$ ($\sigma F \boxdot \varphi \in S$), then for some integer k, $\sigma.k\pi_0 \in S$ ($\sigma.kF\varphi \in S$). S omits infinitely many integers if the set of integers that do not appear in prefixes in S is infinite.

Lindenbaum-Henkin construction is as follows. Enumerate all formulas $\sigma_i\Phi_i$ in the language of GrzA. Then use the following recipe for constructing a maximal consistent set of formulas.

$S_0 = \{1F\varphi\}$;

$$S_{n+1} = \begin{cases} S_n \cup \{\sigma_n\Psi_n\} \text{ if this is consistent and } \Psi_n \text{ is not } F \boxdot \psi_n; \\ S_n \cup \{\sigma_n\pi\} \cup \{\sigma_n.k\pi_0\} \cup \{\sigma_n.kT \boxdot (\psi_n \to \boxdot\psi_n)\} \\ \ \text{if } S_n \cup \{\sigma_n\Psi_n\} \text{ is consistent, } \Psi_n \text{ is a } \pi \text{ formula } (= F \boxdot \psi_n) \\ \quad \text{and } \sigma_n.k \text{ is new} \\ S_n \text{ otherwise.} \end{cases}$$

Here 'new' means that $\sigma_n.k$ does not occur in S_n or in $\sigma_n\pi$. Note that the construction itself is essentially the same as the one for the modal logic K in [59] except that instead of K-consistency we use GrzA-consistency in the construction. We check the following claim concerning this construction and we start from the singleton of $1F\varphi$ (the formula for which we construct a countermodel).

Claim 4.2.4 *S_n omits infinitely many integers.*

Proof For each S_n, there are obviously finitely many sets from S_0. At each step we put at most one new number is used. Thus, we need only finitely many integers (at most n) to construct $S_0, \ldots, S_n$. Since $S_0 = \{1F\varphi\}$, S_0 omits infinitely many integers. Hence, the union of $\{1\}$ from S_0 and the set of the integers used to construct $S_1, \ldots, S_n$ (a finite set of integers) must omit infinitely many integers. ⊠ (claim)

Claim 4.2.5 *If S_n omits infinitely many integers, there will always be a new prefix.*

Proof The proof is similar to the case of S4LPN.

Claim 4.2.6 *If $S_n \cup \{\sigma_n \pi\}$ is GrzA-consistent, then so is $S_n \cup \{\sigma_n\pi, \sigma_n.k\pi_0$, $\sigma_n.kT \boxdot (\psi_n \rightarrow \boxdot\psi_n)\}$, provided that $\sigma_n.k$ is new.*

Proof Suppose $S_n \cup \{\sigma_n\pi\}$ is GrzA-consistent. By definition, this implies that no tableau for $S_n \cup \{\sigma_n\pi\}$ is closed (Note that this set itself is finite). Suppose there is a closed tableau for $S_n \cup \{\sigma_n\pi, \sigma_n.k\pi_0, \sigma_n.kT\boxdot(\psi_n \rightarrow \boxdot\psi_n)\}$. Construct one such tableau. Note that $S_n \cup \{\sigma_n\pi\}$ is a proper subset of this set, and by applying π-rule of Grz on $S_n \cup \{\sigma_n\pi\}$, we will obtain two formulas

$\sigma_n.k\pi_0$ and $\sigma_n.kT \boxdot (\psi_n \to \boxdot\psi_n)$. This implies that the closed tableau we have constructed can be taken to be a closed tableau for the initial set with one additional application of π-rule. (Note that this requires that $\sigma_n.k$ be new.) This implies that we can construct a closed tableau for $S_n \cup \{\sigma_n\pi\}$, which is contradictory to the assumption that $S_n \cup \{\sigma_n\pi\}$ is GrzA-consistent. So, no tableau for $S_n \cup \{\sigma_n\pi, \sigma_n.k\pi_0, \sigma_n.kT \boxdot (\psi_n \to \psi_n\boxdot)\}$ is closed. Hence, $S_n \cup \{\sigma_n\pi, \sigma_n.k\pi_0, \sigma_n.kT \boxdot (\psi_n \to \boxdot\psi_n)\}$ is GrzA-consistent. ⊠ (claim)

Let us define $S_\omega = \bigcup_n S_n$.

Lemma 4.2.7 *If $\{1F\varphi\}$ is GrzA-consistent, then S_ω will be maximally GrzA-consistent and π-complete.*

Proof Suppose $\{1F\varphi\}$ is GrzA-consistent. (Note that this obviously omits infinitely many integers.) We construct S_ω following the construction above. We claim the following two statements.

Claim 4.2.8 *S_ω is GrzA-maximally consistent.*

Proof The proof is similar to the case of S4LPN. ⊠(claim)

Claim 4.2.9 *S_ω is π-complete.*

Proof The proof is similar to the case of S4LPN. ⊠(claim) ⊠(lemma)

Now we construct a canonical Kripke model $\mathcal{K}^s = (K^s, R^s, R^e, r, \mathcal{E}, \Vdash)$ for GrzA based on this maximal GrzA-consistent set. Let K^s be $\{\sigma | \sigma\Phi \in S_\omega\}$.

We identify sequences with possible worlds (so the interpretation function will be the identity function). The accessibility relations R^s and R^e, the root node r, forcing relation $\Vdash$ (for atomic p) and evidence function $\mathcal{E}$ are given as follows:

1. $\sigma R^s \sigma'$ iff σ is a (not necessarily proper) initial segment of σ' ($\sigma \leq \sigma'$).
2. $\sigma R^e \sigma'$ iff $1 \leq \sigma$ and $1 \leq \sigma'$.
3. $r = 1$.
4. $\psi \in \mathcal{E}(\sigma, t)$ iff $\sigma F t : \psi \notin S_\omega$.
5. $\sigma \Vdash p$ iff $\sigma T p \in S_\omega$; $\sigma \Vdash \bot$ iff $\sigma T \bot \in S_\omega$.

Step 2: We have to verify that the R^s, R^e, $\Vdash$ and $\mathcal{E}$ all satisfy the conditions of a Kripke model for GrzA. In this step, we verify all the conditions except one frame condition.

First, it is obvious that $\leq$ is reflexive and transitive. We can show that it is anti-symmetric as follows: suppose $\sigma \leq \sigma'$ and $\sigma' \leq \sigma$. By definition of non-proper initial segment of a sequence of positive integers, this implies the following disjunction ($\sigma < \sigma' \wedge \sigma' < \sigma$) or ($\sigma < \sigma' \wedge \sigma' = \sigma$) or ($\sigma' < \sigma \wedge \sigma = \sigma'$) or ($\sigma = \sigma' \wedge \sigma' = \sigma$). The first one is clearly contradictory. The second one implies $\sigma < \sigma$ and the third one implies $\sigma' < \sigma'$, which are contradictory, too. So, $\sigma = \sigma'$ holds. Second, it easily follows from the definition that $1 \leq \sigma$ and $1 \leq \sigma'$ is an equivalence relation. Also, obviously, $R^s \subseteq R^e$.

We have to further verify the following three things:

1. $\mathcal{E}$ satisfies the conditions of an evidence function;

2. $\Vdash$ can be extended to the entire language of GrzA (Truth Lemma);

3. R^s really satisfies the condition of having no strictly ascending infinite chain.

In this step, for the sake of convenience in giving a presentation of the whole proof in a sensible manner, we show the first two items and leave the complicated condition to the next step.

Proposition 4.2.10

1. *$\sigma Ft : \psi \in S_\omega$ if and only if $\sigma' Ft : \psi \in S_\omega$ (for any σ, σ' in K^s)*
2. *$\sigma Tt : \psi \in S_\omega$ if and only if $\sigma' Tt : \psi \in S_\omega$ (for any σ, σ' in K^s)*

Proof The proof of this proposition is similar to the case of S4LPN.

The following are crucial features of the model that we have constructed and can be proven as corollaries of the above proposition.

Corollary 4.2.11 *Let $\mathcal{K}^s$ be the canonical model we have constructed.*

1. *For all $\sigma_1, \sigma_2 \in K^s$, $\sigma_1 R^e \sigma_2$ (i.e., $R^e = K^s \times K^s$).*
2. *For any $\sigma \in K^s$, $(\sigma \Vdash t : \varphi)$ or for any $\sigma \in K^s$, $(\sigma \nVdash t : \varphi)$.*

Proof 1. By construction, for any $\sigma_1, \sigma_2 \in K^s$, $1 \leq \sigma_1$ and $1 \leq \sigma_2$; hence, $\sigma_1 R^e \sigma_2$.

2. Suppose otherwise, namely $\exists \sigma_1 \in K^s(\varphi \notin \mathcal{E}(\sigma_1, t)$ or $\exists \sigma_2 \in K^s(\sigma_1 R^e \sigma_2$ and $\sigma_2 \nVdash \varphi))$, and $\exists \sigma_3 \in K^s(\varphi \in \mathcal{E}(\sigma_3, t)$ and $\forall \sigma_4 \in K^s(\sigma_3 R^e \sigma_4 \Longrightarrow \sigma_4 \Vdash \varphi))$.

Take an instance for σ_1 and σ_3, respectively, and call them σ'_1 and σ'_3.

The first disjunct of the first statement almost immediately leads to contradiction with the first conjunct of the second statement, due to the above proposition.

Then consider the second disjunct and the last sentence. Take an instance of σ_2 and call it σ'_2. Since $1 \leq \sigma'_1$ and $1 \leq \sigma'_3$, $\sigma'_1 R^e \sigma'_3$. On the other hand, by symmetry, $\sigma'_2 R^e \sigma'_1$ is derived from $\sigma'_1 R^e \sigma'_2$. By transitivity and symmetry, $\sigma'_3 R^e \sigma'_2$. Taking an instantiation of the last universal quantification, $\sigma'_3 R^e \sigma'_2 \Longrightarrow \sigma'_2 \Vdash \varphi$. Hence, $\sigma'_2 \Vdash \varphi$. This contradicts the second disjunct of the first statement $\sigma'_2 \nVdash \varphi$. Therefore, we get a contradiction, in either way. ⊠

We now prove that the constructed $\mathcal{E}$ satisfies the closure conditions of an evidence function.

Proposition 4.2.12 *The evidence function defined above satisfies the following conditions: (1) monotonicity (stability), (2) closure conditions, (3) constant specification.*

Proof The proof is similar to the case of S4LPN. ⊠

Then we can prove the crucial lemma. (Recall that you are taking $\mathcal{N}$ to be identity function.)

Lemma 4.2.13 (Truth Lemma) $\sigma\Psi \in S_\omega \Longrightarrow \Psi$ *is realized at* σ *in the GrzA model* $\mathcal{K}^s$.

Proof By induction on the structure of formulas. The proof is very similar to the case of S4LPN. ⊠ (Truth Lemma)

Step 3: Now let $\mathcal{K}^s = (K^s, R^s, R^e, r, \mathcal{E}, \Vdash)$ be the canonical "model" for GrzA constructed by the Lindenbaum-Henkin construction. But this is not yet a model for GrzA, since does not satisfy the frame condition for a GrzA model, namely that it has no infinite strictly ascending chain. (In this sense, it is still a pseudo-model in the aforementioned terminology). In order to satisfy the condition that a model of GrzA has no infinite strictly ascending chain, we need some care in order to modify the construction so that this model satisfies the condition. We use a method of finitizing the "height" of the model by using the technique of bounded bisimulation explained in [25]. This is the third item listed above.

Before moving on, we give some terminologies about the construction (we follow [25] and [73] in these definitions). The notion of the *height* of states in $\mathcal{K}^s$ is defined by induction. The only element of height 0 is the root of the model; the states of height $n+1$ are those immediate successors of elements of height n that have not yet been assigned a height smaller than $n+1$. The *height of a model* $\mathcal{K}^s$ is the maximum n such that there is a state of height n in $\mathcal{K}^s$, if such a maximum exists; otherwise the height of $\mathcal{K}^s$ is infinite. The model we construct will be such that any strictly ascending chain starting from 1 has a finite length, which satisfies the condition that there is no infinite strictly ascending chain.

Then we introduce the notion of the degree of a modal formula φ.

Definition 4.2.14 *We define the degree of a modal formula as follows.*

1. $deg(p) = 0$*; 2.* $deg(\bot) = 0$*; 3.* $deg(\neg\varphi) = deg(\varphi)$

4. $deg(\varphi * \psi) = max\{deg(\varphi), deg(\psi)\}$*, where* $* \in \{\vee, \wedge. \rightarrow\}$*.*

5. $deg(\Box\varphi) = deg(\varphi) + 1$*.*

6. $deg(t : \varphi) = deg(\varphi) + 1$*.*

We now define n-bisimulation (bounded bisimulation) on arbitrary Fitting models $\mathcal{K} = (K, R, R^e, r, \mathcal{E}, \Vdash)$. This satisfy minimal conditions as follows: there are no conditions for R but we assume $R \subseteq R^e$, $\mathcal{E}$ is stable w.r.t. R^e. [33] The definition of n-bisimulation here is based on [25]. We added some machinery to handle logic of proofs. Let $\mathcal{K}$ and $\mathcal{K}'$ be such Fitting models.

Definition 4.2.15 *We call* w *and* w' *are* n*-bisimilar* ($w \simeq_n w'$) *if there exists a sequence of binary relations* $Z_n \subseteq \cdots \subseteq Z_0$ *with the following properties (for* $i + 1 \leq n$*).*

(0) $\mathcal{E}(v, t) = \mathcal{E}(v', t)$ *for any* $v \in K$ *and for any* $v' \in K'$

(1) wZ_nw'

(2) If vZ_0v'*, then* v *and* v' *agree on all propositional letters.*

(3) If $vZ_{i+1}v'$ *and* vRu*, then* $\exists u'$ $(v'Ru'$ *and* $uZ_iu')$*.*

(4) If $vZ_{i+1}v'$ *and* $v'Ru'$*, then* $\exists u$ $(vRu$ *and* $uZ_iu')$*.*

(5) If $vZ_{i+1}v'$ *and* vR^eu*, then* $\exists u'$ $(v'R^eu'$ *and* $uZ_iu')$*.*

(6) If $vZ_{i+1}v'$ *and* $v'R^eu'$*, then* $\exists u$ $(vR^eu$ *and* $uZ_iu')$*.*

---

[33] In the actually construct model, R^e is universal, so it just matches the underlying idea of an evidence function here, namely evidence situations are the same everywhere. See 4.2.10 and 4.2.11 the statement 1.

The condition (0) is motivated by the stability of evidence function, i.e. $uR^e v$ implies $\mathcal{E}(u,t) = \mathcal{E}(v,t)$ (and in our construction of the model, R^e relates any two world constructed by $F\Box\varphi$ formulas.)

Definition 4.2.16 *We define the notion of a generated subset $U(v)$ of the underlying set K of a model $\mathcal{K}$ from v with respect to R (R^e), respectively, as follows.*

$U(v) = \{u \in K | \exists j(0 \leq j \textit{ and } v(R)^j u)\}$

$U_e(v) = \{u \in K | \exists j(0 \leq j \textit{ and } v(R^e)^j u)\}$

Similarly, we define a bounded generated subset $U^n(v)$ from v with respect to R as follows.

$U^n(v) = \{u \in K | \exists i(0 \leq i \leq n \textit{ and } v(R)^i u)\}$.

In the following argument, we only finitize the height of the canonical model $\mathcal{K}^s$ for GrzA. To do that, we first prove a few general propositions stating properties of bounded bisimulation.[34] By $w \equiv_n w'$, we mean that for any φ s.t. $deg(\varphi) \leq n$, $w \Vdash \varphi$ iff $w' \Vdash \varphi$. We first prove the following lemma.

Lemma 4.2.17 *1. For any $w, v \in K$ and $w', v' \in K'$,*

$w \simeq_{n+1} w'$, wRv, $w'R'v'$, and $vZ_n v'$ implies $v \simeq_n v'$.

2. For any $w, v \in K$ and $w', v' \in K'$,

$w \simeq_{n+1} w'$, $wR^e v$, $w'R^{e\prime}v'$, and $vZ_n v'$ implies $v \simeq_n v'$.

---

[34] The first definition of bisimulation for logic of proofs (justification logic) is given by [136]. Our definition is different from Renne's definition in two ways. 1. It has a bound in depth n. 2. It is simpler, since it uses a strong condition (0) because we have the stability condition for evidence function.

Proof 1. Suppose $w \simeq_{n+1} w'$, wRv, $w'R'v'$, and vZ_nv'.

Then there exists a sequence $Z_{n+1} \subseteq \cdots \subseteq Z_0$ that satisfies the conditions of $w \simeq_{n+1} w'$. We now construct a new sequence of relations (we call them Z'_i $(0 \leq i \leq n)$ based on these Z_i that satisfies the condition of $v \simeq_n v'$.

To do that, take the generated subsets $U_e(v) = \{u \in K | \exists j(0 \leq j$ and $v(R^e)^j u)\}$ and $U'_e(v') = \{u \in K' | \exists j(0 \leq j$ and $v(R^{e\prime})^j u)\}$.[35] $(i = 0$ means that v (or v') itself in the set, respectively.) Let $Z'_i = Z_i \cap (U_e(v) \times U'_e(v'))$ and take the sequence of Z'_i up to n $(0 \leq i \leq n)$, s.t. $Z'_n \subseteq \cdots \subseteq Z'_1 \subseteq Z'_0$. We show that this sequence satisfies all the conditions for $v \simeq_n v'$.

For (0), this condition is the same in $\simeq_{n+1}$ and $\simeq_n$, so it is obvious.

For (1), vZ_nv' is given by assumption, and clearly $(v, v') \in U_e(v) \times U_e(v')$.

For (2), since $Z'_0 \subseteq Z_0$ and Z_0 satisfies the condition (2), Z'_0 satisfies (2).

For (3), suppose $xZ'_{i+1}x'$ and xRy $(1 \leq i + 1 \leq n)$. Since $Z'_{i+1} \subseteq Z_{i+1}$, $xZ_{i+1}x'$ and xRy. By the condition (3) of $w \simeq_{n+1} w'$, $\exists y'(x'R'y'$ and $yZ_iy'))$. Call the state y'_1. xRy and $x'R'y'_1$ implies xR^ey and $x'R^{e\prime}y'_1$. Since $(x, x') \in U_e(v) \times U'_e(v')$ (this follows from $xZ'_{i+1}x'$), $(y, y'_1) \in U_e(v) \times U'_e(v')$. Thus, $yZ'_iy'_1$. Therefore, $\exists y'(x'R'y'$ and $yZ'_iy'))$. For (4), (5) and (6), the proof is similar to the case (3). (Case (5) and (6) are even slightly simpler since we directly consider R^e.)

2. Suppose $w \simeq_{n+1} w'$, wR^ev, $w'R^{e\prime}v'$, and vZ_nv'. The difference from the first case is that in the second case, there is no guarantee that $(w, v) \in R$

---

[35] We use R^e here to take care of both (3), (4) and (5), (6). Using R is not enough to show the conditions (5) or (6).

or $(w', v') \in R'$. (Hence, the difference consists in the case (3) and (4) of 1. The other cases are essentially the same.) However, mainly due to the condition $w \simeq_{n+1} w'$, we can give a similar argument to the case (3) and (4). Indeed, for case (3), suppose $xZ'_{i+1}x'$ and xRy $(1 \leq i+1 \leq n)$. By definition, $xZ'_{i+1}x'$ implies $xZ_{i+1}x'$ and $(x, x') \in U_e(v) \times U_e(v')$. On the other hand, by $w \simeq_{n+1} w'$, if $xZ_{i+1}x'$ and xRy, then $\exists y'(x'Ry'$ and $yZ_iy')$. Hence, $\exists y'(x'Ry'$ and $yZ_iy')$. Call it y'_1. $x'Ry'_1$ and $yZ_iy'_1$. Since xRy, $x'Ry'_1$, and $(x, x') \in U_e(v) \times U_e(v')$, $(y, y'_1) \in U_e(v) \times U_e(v')$. Thus, $(y, y'_1) \in Z_i \cap (U_e(v) \times U_e(v'))$. Hence, $yZ'_iy'_1$. Therefore, $\exists y'(x'Ry'$ and $yZ'_iy')$ The other cases are the same. ⊠

Now we are ready to state and prove the desired property of n-bisimulation.

Proposition 4.2.18 *For any $w \in \mathcal{K}$ and $w' \in \mathcal{K}'$, $w \simeq_n w' \implies w \equiv_n w'$.*

Proof Proof by induction on n. Base case: $n = 0$. $deg(\varphi) = 0$. Here φ is either a propositional variable and their Boolean combination. Due to the condition (1) and (2) for $w \simeq_0 w'$, i.e. wZ_0w', and if wZw' then $w \Vdash p$ iff $w' \Vdash p$. Boolean cases are straightforward.

Inductive case: We assume the statement for n, namely $\mathcal{K}, w \simeq_n \mathcal{K}', w'$ $\implies w \equiv_n w'$, and prove the statement for $n+1$, namely $\mathcal{K}, w \simeq_{n+1} \mathcal{K}', w'$ $\implies w \equiv_{n+1} w'$.

We have two subcases.

Subcase 1. $\varphi = \Box\psi$ $(deg(\varphi) = deg(\Box\psi) = \text{deg}(\psi) + 1 = n + 1)$. We want to show that $w \simeq_{n+1} w' \implies w \nVdash \Box\psi$ iff $w' \nVdash \Box\psi$.

To show this, suppose (A) $w \simeq_{n+1} w'$, and suppose (B) $w \nVdash \boxdot\psi$, i.e. $\exists v(wRv$ and $v \nVdash \psi)$. Call this state v_1. Then, wRv_1 and $v_1 \nVdash \psi$. By the condition (1) of the assumption (A), $wZ_{n+1}w'$. So, we have $wZ_{n+1}w'$ and wRv_1. By condition (3) of the assumption (A), $\exists v'(w'R'v'$ and $v_1Z_nv')$. Call this v'_1. So $w'R'v'_1$ and $v_1Z_nv'_1$.

Note that we now have all the statements in the assumptions of the lemma with respect to particular instances v_1, v'_1. Hence, $v_1 \simeq_n v'_1$.

By IH, $v_1 \simeq_n v'_1 \Longrightarrow v_1 \Vdash \psi$ iff $v'_1 \Vdash \psi$. Since the antecedent of the claim is already shown above, we have $v'_1 \nVdash \psi$. So, $\exists v' \in K(w'R'v'$ and $v' \nVdash \psi)$. Therefore, $w' \nVdash \boxdot\psi$. The other direction is similar.

Subcase 2. $\varphi = t : \psi$ $(deg(\varphi) = deg(t : \psi) = deg(\psi) + 1)$.

We want to show that $w \simeq_{n+1} w' \Longrightarrow w \nVdash t : \psi$ iff $w' \nVdash t : \psi$. To show this, suppose (A) $w \simeq_{n+1} w'$, and suppose (B) $w \nVdash t : \psi$, i.e. $\psi \notin \mathcal{E}(w, t)$ or $\exists v \in K(wR^e v$ and $v \nVdash \psi)$. (Call it v_1.)

By the condition (0) of $w \simeq_{n+1} w'$, without depending the number $n + 1$, the first disjunct implies $\psi \notin \mathcal{E}'(w', t)$. So, $\psi \notin \mathcal{E}'(w', t)$ or $\exists v' \in K'(wR^{e\prime}v'$ and $v' \nVdash \psi)$ is implied by the first disjunct.

By the condition (1) of the assumption (A), $wZ_{n+1}w'$. Also, the second disjunct of (B) implies $wR^e v_1$. Hence, by the condition (5) of (A), $\exists v' \in K'(v_1Z_nv'$ and $w'R^{e\prime}v')$ (Call it v'_1). So, we have $w'R^{e\prime}v'_1$ and $v_1Z_nv'_1$.

By the above lemma (the statement 2), since we have shown the particular instance of the assumptions of this lemma, we can show $v_1 \simeq_n v'$.

By IH, $v_1 \simeq_n v'_1 \Longrightarrow v_1 \nVdash \psi$ iff $v'_1 \nVdash \psi$. Hence, $v_1 \nVdash \psi$ iff $v'_1 \nVdash \psi$.

Thus, $v'_1 \nVdash \psi$. So, we have $\exists v' \in K'(w'R^{e\prime}v'$ and $v' \nVdash \psi)$.

So, either way, $\psi \notin \mathcal{E}'(w', t)$ or $\exists v' \in K'(wR^{e\prime}v'$ and $v' \nVdash \psi)$ is derivable. Therefore, $w' \nVdash t : \psi$. The converse is similar. ⊠ (proposition)

Step 4: The countermodel to be constructed has to satisfy the required frame condition for the underlying frame of a model of GrzA, i.e. there is no infinite strictly ascending chain. Thus, we now construct a finite-height model that is n-bisimilar to the original canonical model $\mathcal{K}^s$ for GrzA. Before stating the proposition, we give a few definitions. We use the notation $\mathcal{K}[r]$ to stand for a model $\mathcal{K}$ with the root node r. Also, let $U^n(r) = \{v \in K^s | \exists i(0 \leq i \leq n$ and $r(R^s)^i v)\}$, taking r in the definition of a bounded generated subset of K^s. (Clearly, $U^n(r) \subseteq K^s$, and $U^n(r) \times U^n(r) \subseteq K^s \times K^s$.) The restriction of $\mathcal{K}^s$ to $U^n(r)$ (we use the notation $\mathcal{K}^s \restriction U^n(r)$) is defined as follows. The underlying set of $\mathcal{K}^s \restriction U^n(r)$ is $K^s \cap U^n(r)$, accessibility relations are restriction of R^s, R^e, i.e. $R^s \cap (U^n(r) \times U^n(r))$ (we use the notation R^{s^*} for this) and $R^e \cap (U^n(r) \times U^n(r))$ (we use the notation R^{e^*} for this), the evidence function is a restriction of the function w.r.t. the domain of the function $\mathcal{E} \restriction U^n(r)$ (we use the notation $\mathcal{E}^*$ for this), and the forcing relation is also a restriction of the first coordinate of the relation, $\Vdash \restriction U^n(r)$ (we use the notation $\Vdash^*$ for this). Hence, we write $\mathcal{K}^s \restriction U^n(r) = (K^s \cap U^n(r), R^{s^*}, R^{e^*}, r, \mathcal{E}^*, \Vdash^*)$. The construction of n-bisimulation defined above can be applied to these two models $\mathcal{K}^s \restriction U^n(r)$ and $\mathcal{K}^s$.

Proposition 4.2.19 *$(\mathcal{K}^s \restriction U^n(r), r) \simeq_n (\mathcal{K}^s, r)$ for each $n \in \mathbb{N}$.*

In order to define n-bisimulation, it suffices to use $n+1$-sequence of relations $Z_n \subseteq Z_{n-1} \subseteq \cdots \subseteq Z_0$ such that the conditions of bounded bisimulation (3), (4), (5) and (6) hold among R^s, R^{s^*}, R^e, R^{e^*} and Z_i for any $0 \leq i \leq n$. We want to show that a rooted model $\mathcal{K}^s[r]$ restricted to $U^n(r)$ is sufficient to show that n-bisimulation always holds between $(\mathcal{K}^s, r)$ and $(\mathcal{K}^s[r] \restriction U^n(r), r)$.

Proof Proof is by induction on n.

Base case: $n = 0$. (0) is obvious, since one is a submodel of the other and $\mathcal{E}$ coincides everywhere. (1) is rZ_0r (take the root r). This follows since each of these is the only node of each model. (2) $vZ_0v' \Rightarrow v \Vdash p$ iff $v' \Vdash^* p$. This holds since again we are taking the roots for both cases (if $n = 0$, then $v = r$ and $v' = r$). (3), (4), (5), and (6) are vacuously correct since we do not have either uRv or $uR'v'$ (except reflexivity since we are discussing the case of GrzA) due to the fact that $n = 0$.

Inductive case: By IH, suppose the statement holds for k, i.e. $(\mathcal{K}^s[r] \restriction U^k(r), r) \simeq_k (\mathcal{K}^s, r)$. We show that $(\mathcal{K}^s[r] \restriction U^{k+1}(r), r) \simeq_{k+1} (\mathcal{K}^s, r)$. Given $Z_k \subseteq \cdots \subseteq Z_0$ ($n = k$) for $\simeq_k$, we define $Z'_{k+1}, \ldots, Z'_0$ for $n = k+1$ as follows.

Let $Z'_{k+1} := \{(r, r)\}$. $Z'_i := Z_i \cup \{(v, v) | r(R^s)^{k+1-i}v\}$ $(0 \leq i \leq k)$.

(0) is obvious for the same reason as in the base case.

(1) By definition, $(r, r) \in Z'_{k+1}$.

(2) For $\mathcal{K}$ and the submodel of $\mathcal{K}$ up to the height $k+1$, propositional letters obviously agree in both of them.

(3) All the cases of Z'_i s.t. $1 \leq i \leq k+1$ come from induction hypothesis,

since Z'_i are the same as Z_{i-1} in these cases. Since $Z'_1 = Z_1 \cup \{(v, v)|r(R^s)^k v\}$, $Z'_0 = Z_0 \cup \{(v, v)|r(R^s)^{k+1} v\}$, which goes beyond the $\simeq_k$. So the only case that should be verified is the case where $i = 0$.

Pick any state w s.t. $r(R^s)^k w$. Suppose $wZ'_1 w$ and $wR^s v$. Since $r(R^s)^k w$, $r(R^s)^{k+1} v$ holds. Hence, obviously $(w, v) \in U^{k+1}(r) \times U^{k+1}(r)$. So, $(w, v) \in R^s \cap (U^{k+1}(r) \times U^{k+1}(r))$. Also, since $Z'_0 = Z_0 \cup \{(v, v)|r(R^s)^{k+1} v\}$, $Z'_1 = Z_0$ and $vZ'_0 v$. Hence, $\exists v'[vZ'_0 v'$ and $(w, v') \in R^s \cap (U^{k+1}(r) \times U^{k+1}(r))]$,

(4) Similarly, we check only $i = 0$. Again, we consider any state w s.t. $r(R^s)^k w$. Suppose $wZ'_1 w$ and $(w, v) \in R^s \cap (U^{k+1}(r) \times U^{k+1}(r))$. Then, by definition of Z_0 and since $R^s \cap (U^{k+1}(r) \times U^{k+1}(r)) \subseteq R^s$, clearly $\exists v'$ s.t. $vZ_0 v'$ and $wR^s v'$.

(5) and (6) are similar. Hence, all the conditions are satisfied for $k + 1$ case. Therefore, $(\mathcal{K}^s[r] \upharpoonright U^{k+1}(r), r) \simeq_{k+1} (\mathcal{K}^s, r)$ holds. ⊠

Step 5: This suffices to show that for any formula φ s.t. $deg(\varphi) = n$, $\mathcal{K}^s, r \Vdash \varphi$ iff $\mathcal{K}^s[r] \upharpoonright U^n(r), r \Vdash^* \varphi$. The height of the model $\mathcal{K}^s[r] \upharpoonright U^n(r)$ is n. Hence, for a formula φ for which we have constructed the canonical model $\mathcal{K}^s$ s.t. $\mathcal{K}^s, 1 \nVdash \varphi$, we have a model $\mathcal{K}^s[r] \upharpoonright U^n(r)$, such that $\mathcal{K}^s[1] \upharpoonright U^n(1), 1 \nVdash^* \varphi$, where $r = 1$ and the height of $\mathcal{K}^s[1] \upharpoonright U^n(1)$ is at most $deg(\varphi) = n$. This model may still be an infinite model since there may be some infinite branching in the tree. However, for our purpose, a finite-height model suffices.

Proposition 4.2.20 *1.* R^{s^*} *has no infinite strictly ascending chain.*

2. R^{e^*} *is an equivalence relation. Moreover, for any* $\sigma, \sigma' \in K^s \cap U^n(1)$, $\sigma R^{e^*} \sigma'$*.*

Proof The set of sequences K^s is a set of finite sequences of positive integers that constitute a tree partial order. But sequences have no upperbound in their length, so the set of sequences can have an infinite strictly ascending chain. However, $\mathcal{K}^s \upharpoonright U^n(r)$ has an upperbound in the lengths of sequences that are in the set as elements, so there is no set of sequences in which there is no upperbound in the length of sequences. So, there is no infinite strictly ascending chain.

In the construction of $\mathcal{K}^s[r] \upharpoonright U^n(r)$, we start from the particular canonical model constructed by Lindenbaum-Henkin construction, and we have cut up the irrelevant parts of the model. This keeps R^{e^*} as an equivalence relation, for in our construction, the equivalence relation R^{e^*} is defined as $1 \leq \sigma$ and $1 \leq \sigma'$ (intuitively, two sequences that represent two worlds share the initial segment 1 (the root node)). Since we keep the root node of the original infinite tree partial order, the equivalence relation in the original model was actually a universal relation and the equivalence relation in the cut-up model is also a universal relation in it. Concerning R^e, we simply take a universal relation in $\mathcal{K}(r) \upharpoonright U^n(1)$. ⊠ (proposition)

Hence, the following proposition follows.

Proposition 4.2.21 *For any* ψ*, s.t.* $deg(\psi) = n$, $\mathcal{K}^s[1] \upharpoonright U^n(1), 1 \Vdash^* \psi$ *iff*

$\mathcal{K}^s, 1 \Vdash \psi$.

Combined with the Lindenbaum-Henkin construction, this implies the following theorem.

Theorem 4.2.22 (Weak Completeness) *If* $\{1F\varphi\}$ *has no closed tableau, then there exists a finite-height* GrzA *model* $\mathcal{K}$*, s.t.* $\mathcal{K} = (K, R, R^e, r, \mathcal{E}, \Vdash)$*, s.t.* $r \nVdash \varphi$*.*

Proof Suppose φ is not provable using the prefixed GrzA-tableau rules. Then $\{1F\varphi\}$ is GrzA-consistent, and it omits infinitely many integers. So, we can extend it to a (restricted) maximally GrzA-consistent, π-complete set S_ω by the above construction. We can define a canonical Kripke model $\mathcal{K}^s$ out of S_ω. By Truth Lemma, we can show $F\varphi$ is realized at 1 in $\mathcal{K}^s$. $\mathcal{K}^s, 1 \nVdash \varphi$. By finitizing the height of the model using the proposition presented above, this is equivalent to $\mathcal{K}^s[1] \upharpoonright U^n(1), 1 \nVdash^* \varphi$. Hence, there is a finite height Kripke model $\mathcal{K}$ and the root r such that $\mathcal{K}, r \nVdash \varphi$. ⊠

4.2.2.2 Prefixed Tableau System for wGLA

Now we go back to GLA. Recall we defined an auxiliary subsystem wGLA of GLA without the axiom 3)-3 $t{:}\Box\varphi \to \varphi$ and Reflection rule, 3. $\dfrac{\Box\varphi}{\varphi}$. Both of them are sound with respect to Kripke semantics with the appropriate root soundness condition. However, directly proving completeness with respect to models satisfying the root soundness condition looks difficult, so we first give a prefixed tableau system for wGLA and prove completeness of wGLA with respect to the class of GLA models. (Recall that we have decided to call a model for wGLA a "GLA model.") Later, we consider what rules should be added in order to formulate a prefixed tableau system for the entire GLA. Since we have to introduce an auxiliary prefixed tableau system to eventually handle completeness for GLA, we introduce an abbreviation for naming a prefixed tableau system. We call the prefixed tableau system for wGLA "TwGLA."

In the following, we keep using Fitting's terminology for basic notions in the prefixed tableau system. The only difference from GrzA is this: σ' is accessible from σ iff $\sigma < \sigma'$ ($<$ means that "is a proper initial segment of").

α-rule and β-rules are the same as those of TGrzA.

Rules for LP: These rules are the same as those of TGrzA

Modal Rules:

ν-rules: K $\dfrac{\sigma T \Box \varphi}{\sigma.n T \varphi}$ ($\sigma.n$ is used.) 4 $\dfrac{\sigma T \Box \varphi}{\sigma.n T \Box \varphi}$ ($\sigma.n$ is used.)

The π-rule for GL: $\dfrac{\sigma F \Box \varphi}{\begin{array}{l}\sigma.n F \varphi\\ \sigma.n T \Box \varphi\end{array}}$ ($\sigma.n$ is new.)

In addition, we have Constant Specification Rules as follows: a branch is closed if it has $\sigma Fc : A$ (A is an axiom of wGLA and $A \in \mathcal{CS}(c)$.)

Let us give a definition of an interpretation function for wGLA. A (partial) function $\mathcal{N}$ from the set of σ-sequences to the set of worlds K is a wGLA interpretation (partial) function if, for any $\sigma, \sigma' \in Dom(\mathcal{N})$,

(1) $\exists u$ $(\mathcal{N}(\sigma)Ru)$ and σ' is accessible from $\sigma \Longrightarrow \mathcal{N}(\sigma)R\mathcal{N}(\sigma')$ and

(2) $\exists u$ $(\mathcal{N}(\sigma)R^e u)$ and σ' is e-accessible from $\sigma \Longrightarrow \mathcal{N}(\sigma)R^e\mathcal{N}(\sigma')$.

We consider only $\mathcal{N}$ s.t. $\mathcal{N}(1) = r$. Again, concerning an interpretation realizing a prefixed signed formula, etc., we follow Fitting's terminology (mainly, in [57]) as we did for GrzA. Note that we use the same convention for a prefix being "used" on a branch as we used in GrzA. We prove soundness of prefixed tableau system wGLA with respect to the class of GLA models.

Lemma 4.2.23 *Suppose $\mathcal{T}$ is a satisfiable tableau. If any tableau rule for wGLA is applied to $\mathcal{T}$, then the resulting tableau is still satisfiable.*

Proof The proof is similar to GrzA except π-rule for GL. Since this rule for prefixed tableau has not been present in the literature, it may be worth showing.

π-rule for $\Box$: Suppose (1) $\sigma F\Box\varphi$ in θ and θ is satisfiable, but (2) $\theta \cup \{\sigma.nF\varphi,\ \sigma.nT\Box\varphi\}$ is not satisfiable. By the definition of satisfiability (since $\exists\mathcal{K}\exists\mathcal{N}$ s.t. for any $\sigma_i T\psi_i \in \theta$, $\mathcal{K}, \mathcal{N}(\sigma_i) \Vdash \psi_i$ and for any $\sigma_i F\psi_i \in \theta$, $\mathcal{K}, \mathcal{N}(\sigma_i) \nVdash \psi_i$) and for some $\in \theta$, ψ_j, $\psi_j = F\Box\varphi$), (1) implies $\exists\mathcal{K}\exists\mathcal{N}$, s.t.

$\mathcal{K}, \mathcal{N}(\sigma) \nVdash \Box\varphi$. (2) implies $\forall\mathcal{K}\forall\mathcal{N}$, if, for any $\sigma_i T\psi_i \in \theta$, $\mathcal{K}, \mathcal{N}(\sigma_i) \Vdash \psi_i$ and for any $\sigma_i F\psi_i \in \theta$, $\mathcal{K}, \mathcal{N}(\sigma_i) \nVdash \psi_i$), then it is not the case that $\mathcal{N}(\sigma.n) \nVdash \varphi$ and $\mathcal{N}(\sigma.n) \Vdash \Box\varphi$. Take a particular instance of $\mathcal{K}$ and $\mathcal{N}$, i.e., $\mathcal{K}_1$ and $\mathcal{N}_1$. $\mathcal{K}_1, \mathcal{N}_1(\sigma) \nVdash \Box\varphi$. (Let $\mathcal{N}_1(\sigma) = t$.) So, $\exists v$ ($\mathcal{N}_1(\sigma)Rv$ and $v \nVdash \varphi$). Call this state v_1. So, $\mathcal{N}_1(\sigma)Rv_1$ and $v_1 \nVdash \varphi$.

We extend $\mathcal{N}_1$ so that $\mathcal{N}_1(\sigma.n) = v_1$. So, $\mathcal{N}_1(\sigma)R\mathcal{N}_1(\sigma.n)$ and $\mathcal{N}_1(\sigma.n) \nVdash \varphi$.

[We have to check that the extended $\mathcal{N}_1$ is still an wGLA interpretation. For R, we do not have the case of reflexivity. Since R is irreflexive, $\neg v_1 R v_1$, $\neg\mathcal{N}(\sigma.n)R\mathcal{N}(\sigma.n)$. Otherwise, the proof is similar to GrzA. For R^e, the proof is the same as that of GrzA.]

(2) implies that if $\mathcal{K}_1$, $\mathcal{N}_1(\sigma.n) \Vdash \Box\varphi$, then $\mathcal{K}_1, \mathcal{N}_1(\sigma.n) \Vdash \varphi$ by taking an instance of $\mathcal{K}$ and $\mathcal{N}$ as $\mathcal{K}_1$ and (extended) $\mathcal{N}_1$. So, suppressing $\mathcal{K}_1$, we have that if $\mathcal{N}_1(\sigma.n) \Vdash \Box\varphi$, then $\mathcal{N}_1(\sigma.n) \Vdash \varphi$. Hence, $\mathcal{N}_1(\sigma.n) \nVdash \Box\varphi$, i.e. $v_1 \nVdash \Box\varphi$. So, $\exists u$ (vRu and $u \nVdash \varphi$). (Call this u_1.)

We now consider another interpretation $\mathcal{N}_2$ such that (i) for any prefixed formula $\sigma\Phi \in \theta$, $\mathcal{N}_2(\sigma) = \mathcal{N}_1(\sigma)$ and (ii) $\mathcal{N}_2(\sigma.n) = u_1$. (In $\mathcal{N}_2$, u_1 takes the role that v_1 took in $\mathcal{N}_1$. Thus, we can apply the same argument to u_1 as used to generate a new world u_1 accessible from v_1 by using the same interpretation.)

Take an instance of (2) using this interpretation. Then, $\mathcal{N}_2(\sigma.n) \Vdash \Box\varphi$ implies $\mathcal{N}_2(\sigma.n) \Vdash \varphi$. It follows that $u_1 \Vdash \Box\varphi$ implies $u_1 \Vdash \varphi$. However, since $u_1 \nVdash \varphi$, $u_1 \nVdash \Box\varphi$. Hence, we can get $\exists u_2$ ($u_1 R u_2$ and $u_2 \nVdash \varphi$). (Also, this

u_2 cannot be identical with either t, v_1 or u_1. If u_2 were identical with one of these, that would imply that u_2Ru_2 by transitivity. Contradictory to the frame condition of GL. Hence $u_2 \neq \mathcal{N}(\sigma)$ and $u_2 \neq v_1$ and $u_2 \neq u_1$.)

By the same argument, we can go further, $\exists u_3$ (u_2Ru_3 and $u_3 \nVdash \varphi$), and $u_3 \neq t$, $u_3 \neq v_1$, $u_3 \neq u_2$, etc. Clearly, this argument can be repeated indefinitely, but we will then produce $tRv_1Ru_1Ru_2Ru_3\ldots$ (and this cannot go into a circle since if it happens, say u_1Ru_n and u_nRu_1, then by transitivity we get u_nRu_n, which is impossible), but this contradicts the feature of a GLA model that there is no infinite ascending chain of R in it. Therefore, if θ is satisfiable and $\sigma F\Box\varphi \in \theta$, then $\theta \cup \{\sigma.nF\varphi, \sigma.nT\Box\varphi\}$ is satisfiable. This is sufficient for proving the case of π-rule for wGLA. ⊠(Lemma)

Now we can state and prove the soundness theorem.

Theorem 4.2.24 (Soundness for wGLA) *If φ has a prefixed wGLA-tableau proof, then φ is valid in all GLA models.*

Proof Similar to the cases of S4LPN and GrzA. ⊠

Then we prove the completeness theorem by Lindenbaum-Henkin construction ([59]). The construction is very similar to the case of GrzA. We modify the case in which $\Psi_n = F\Box\psi_n$, which is a π formula. Here is the crucial part of the construction. First, enumerate all prefixed signed formulas $\sigma_1\Psi_1, \ldots, \sigma_n\Psi_n, \ldots$

$S_0 = \{1F\varphi\}$;

$$S_{n+1} = \begin{cases} S_n \cup \{\sigma_n\Psi_n\} \text{ if this is consistent and } \Psi_n \text{ is not } F\Box\psi_n; \\ S_n \cup \{\sigma_n\Psi_n\} \cup \{\sigma_n.k\pi_0\} \cup \{\sigma_n.kT\Box\psi_n\} \\ \quad \text{if } S_n \cup \{\sigma_n\Psi_n\} \text{ is consistent, } \Psi_n \text{ is a } \pi \text{ formula } (= F\Box\psi_n) \\ \qquad \text{and } \sigma_n.k \text{ is new} \\ S_n \text{ otherwise.} \end{cases}$$

We construct a maximal consistent and π-complete set S_ω. The details of the proof here is similar to the case of S4LPN and GrzA, so we omit them here.

Based on this S_ω, we define a canonical GLA pseudo-model $\mathcal{K} = (K, R, R^e, r, \mathcal{E}, \Vdash)$. Let K be $\{\sigma | \sigma\Phi \in S_\omega\}$. We identify sequences with possible worlds (so the interpretation function will be the identity function). The accessibility relations R and R^e, the forcing relation $\Vdash$ for atomic formulas, and the evidence function $\mathcal{E}$ are given as follows:

1. $\sigma R \sigma'$ iff σ is a proper initial segment of σ' ($\sigma < \sigma'$).
2. $\sigma R^e \sigma'$ iff $1 \leq \sigma$ and $1 \leq \sigma'$.
3. $\sigma \Vdash p$ iff $\sigma Tp \in S_\omega$; $\sigma \Vdash \bot$ iff $\sigma T\bot \in S_\omega$.
4. $\psi \in \mathcal{E}(\sigma, t)$ iff $\sigma Ft : \psi \notin S_\omega$.
5. $r = 1$

We have to verify that the R, R^e, $\Vdash$ and $\mathcal{E}$ all satisfy the conditions of a

GLA model. It is obvious by definition that $<$ is transitive, and $1 \leq \sigma$ and $1 \leq \sigma'$ is an equivalence relation. Also, obviously, $R \subseteq R^e$. We have to prove that R really satisfies the condition that it has no infinite ascending chain.

We first show that the conditions for $\mathcal{E}$ are satisfied and that the forcing relation can be extended to the entire object language (Truth Lemma). Since the proof is similar to the cases of S4LPN and GrzA, we omit the details here.

Proposition 4.2.25

1. *$\sigma Ft : \psi \in S_\omega$ if and only if $\sigma' Ft : \psi \in S_\omega$ (for any $\sigma, \sigma' \in K$)*
2. *$\sigma Tt : \psi \in S_\omega$ if and only if $\sigma' Tt : \psi \in S_\omega$ (for any $\sigma, \sigma' \in K$)*

Proposition 4.2.26 *The evidence function defined above satisfies the following conditions: (1) monotonicity, (2) closure conditions, (3) constant specification.*

Lemma 4.2.27 (Truth Lemma)

$\sigma\Psi \in S_\omega \Longrightarrow \Psi$ is realized at σ in a GLA model $\mathcal{K}$.

Proof Similar to the cases of S4LPN and GrzA. (In one place of the case $\Box\varphi$, we have to replace the use of ν_T by ν_K by taking the initial segment (one-element shorter) of the relevant prefix.) ⊠

As we discussed in the case of GrzA, our canonical pseudo-model construction itself does not guarantee that the constructed pseudo-model satisfies the condition of having no infinite ascending chain. Hence, we finitize the height of the canonical pseudo-model $\mathcal{K}$ by the same construction as we used for GrzA, i.e. by bounded bisimulation. Since the construction of bounded bisimulation used for GrzA does not depend on reflexivity of the accessibility relation in the model, we can construct a model $\mathcal{K}^*$ of wGLA whose height is finite by using the same technique as the case of GrzA. Thus, we omit the details here. We have the following proposition.

Proposition 4.2.28 *R^* constructed above is a converse well-founded relation on K^*.*

Proof Similar to the case of GrzA. ⊠

Theorem 4.2.29 (Weak Completeness for wGLA)

If $\{1F\varphi\}$ has no closed tableau, then there exists a GLA Kripke model $\mathcal{K}^*$ $(= (K^*, R^*, R^{e^*}, r^*, \mathcal{E}^*, \Vdash^*))$, s.t. $1 \nVdash^* \varphi$.

Proof Similar to the case of GrzA. ⊠

4.2.2.3 Prefixed tableau system for GLA

Now we extend the completeness theorem from wGLA to the full GLA. We first formulate a prefixed tableau system for GLA, which we call TGLA. It turns out that the axiom 3)-3 $t : \Box\varphi \to \varphi$ and Reflection Rule $\dfrac{\Box\varphi}{\varphi}$ can be dealt with by adding one tableau rule called "Reflection Rule." We need to add Reflection Rule and the following modification of TwGLA to obtain

TGLA.

1. Constant specification has to be modified accordingly. Instead of taking all the axioms for wGLA, we use all the axioms of the full GLA.

2. Reflection Rule: From $\{1F\varphi\}$, $\dfrac{\cancel{1F\varphi}}{\{1F\Box\varphi\}}$ can be derived.

Note : This crossing is to emphasize that $1F\varphi$ is not on the same tableau after an application Reflection Rule, but the result of an application of the rule should be taken as a tableau derivation of a formula below the line from the formula on the line. The crucial feature of this rule that makes it different from more conventional rule in prefixed tableau systems is that although satisfiability is preserved from the premise to the conclusion, we have moved from one model $\mathcal{K}$ (in the premise) to another $\mathcal{K}^+$ (in the conclusion). We will discuss the details of these models later. Accordingly, the premise and the conclusion have the same prefix 1, but their interpretations of the prefix 1 are different. Not retaining the premise after an application of Reflection Rule can be regarded as a kind of destructive feature of this rule. This destructive feature of Reflection Rule is useful to translate the tableau system into a hypersequent calculus.

In addition to these modifications in our basic formulation of TGLA from TwGLA, we have to reformulate what we take to be a proof of the tableau system TGLA.

Definition 4.2.30 *φ has a tableau proof in TGLA if $\{1F\Box\varphi\}$ has a closed tableau in TGLA.*

An example of a tableau proof : $t : \Box\Box\varphi \to \varphi$ is derivable by this rule.[36]

1. ~~$1F\Box(t : \Box\Box\varphi \to \varphi)$~~
2. $1F\Box\Box(t : \Box\Box\varphi \to \varphi)$ (Reflection Rule, line 1)
3. $1.1F\Box(t : \Box\Box\varphi \to \varphi)$; $1.1T\Box\Box(t : \Box\Box\varphi \to \varphi)$ (π-rule, line 2)
4. $1.1.1F(t : \Box\Box\varphi \to \varphi)$; $1.1.1T\Box(t : \Box\Box\varphi \to \varphi)$ (π-rule, line 3)
5. $1.1.1Tt : \Box\Box\varphi$ (α-rule, line 4)
6. $1.1.1F\varphi$ (α-rule, line 4)
7. $1.1Tt : \Box\Box\varphi$ (E4r, line 5)
8. $1Tt : \Box\Box\varphi$ (E4r, line 7)
9. $1T\Box\Box\varphi$ (ET, line 8)
10. $1.1T\Box\varphi$ (ν-rule for K, line 9)
11. $1.1.1T\varphi$ (ν-rule for K, line 10)

$\times$

Note : This naturally shows that although we use an additional $\Box$ in the definition of a tableau proof in TGLA, Reflection Rule is not redundant due to the possibility of its iterated applications.

---

[36] To save space, we use ";" to write more than one formula horizontally.

In the following, we first show soundness of TGLA. It suffices to prove soundness of Reflection Rule since soundness of all other rules are already shown.

If we consider the meaning of Reflection Rule from a semantic point of view, then it would be clear that this rule corresponds to the operation of gluing another root below a given model. We show the soundness of reflection rule with respect to the appropriate root sound Kripke semantics that we have already defined.

Lemma 4.2.31 *Suppose $\mathcal{T} = \{1F\varphi\}$ is a tableau that is satisfiable in a φ-sound GLA model. If reflection rule for GLA is applied to $\mathcal{T}$, then the resulting tableau $\{1F\Box\varphi\}$ is still satisfiable in a $\Box\varphi$-sound GLA model.*

Proof Suppose $\{1F\varphi\}$ is satisfiable in a φ-sound GLA model. We want to show that $\{1F\Box\varphi\}$ is satisfiable in a $\Box\varphi$-sound GLA model. (Note that $\{\Box\psi \to \psi | \Box\psi \in Sb(\varphi)\} \subseteq \{\Box\psi \to \psi | \Box\psi \in Sb(\Box\varphi)\}$. In addition, the only formula from the latter set that is missing from the former is $\Box\varphi \to \varphi$, since $Sb(\Box\varphi) \backslash Sb(\varphi) = \{\Box\varphi\}$.)

Satisfiability implies that there exist an GLA-interpretation (partial) function $\mathcal{N}$ and a GLA model that is φ-sound, s.t. $\mathcal{K}, \mathcal{N}(1) \nVdash \varphi$ and $\mathcal{N}(1) = r$ ($r \in K$). To show $1F\Box\varphi$ is satisfiable in a $\Box\varphi$-sound GLA model, we construct another GLA-model that is $\Box\varphi$-sound and an interpretation function $\mathcal{N}^+$ that satisfies $\{1F\Box\varphi\}$, following the argument gluing a new root node given in [4]. Let $\mathcal{K}^+$ be a sextuple $(K^+, R^+, R^{e^+}, r^+, \mathcal{E}^+, \Vdash^+)$, s.t.

1. $K^+ = \{r^+\} \cup K$, where r^+ is a new root node of $\mathcal{K}^+$ ($r^+ \notin K$).

2. $R^+ = \{(r^+, r)\} \cup \{(r^+, y) | (r, y) \in R\} \cup R$

3. $R^{e^+} = \{(r^+, y) | y \in K^+\} \cup R^e$.

4. $\mathcal{E}^+$ is s.t. $\forall u \in K$ $[\mathcal{E}^+(r^+, t) = \mathcal{E}(u, t)$ and $\mathcal{E}^+(u, t) = \mathcal{E}(u, t)]$.[37]

5. The forcing relation $\Vdash^+$ is defined as follows.

5.1. For any propositional variable $p \in Sb(\varphi)$,

if $u = r^+$, then $u \Vdash^+ p$ iff $r \Vdash p$, and if $u \neq r^+$, then $u \Vdash^+ p$ iff $u \Vdash p$.

5.2. For all $u \in K^+$, $r^+ \nVdash^+ \bot$.

5.3. At any $u \in K^+$, $\Vdash^+$ commutes with Booleans for any formula in $Sb(\varphi)$.[38]

5.4. For any $\Box\psi \in Sb(\varphi)$, $\forall u \in K^+$, $u \Vdash^+ \Box\psi$ iff $\forall v \in K^+$, $uR^+v \Longrightarrow v \Vdash^+ \psi$

5.5. For any $t : \psi \in Sb(\varphi)$, $\forall u \in K^+$, $u \Vdash^+ t : \psi$ iff $\psi \in \mathcal{E}^+(u, t)$ and $\forall v \in K^+$ $(uR^{e^+}v \Longrightarrow v \Vdash^+ \psi)$

The following are additional conditions: 6. (Constant Specification)

$A \in \mathcal{CS}(c)$ (($\mathcal{CS}$) for GLA) implies $K^+, u \Vdash^+ c : A$ for all $u \in K^+$

---

[37] The definition of $\mathcal{E}^+$ goes beyond $Sb(\Box\varphi)$ due to its closure conditions. However, in order to prove this lemma, we do not have to go beyond $Sb(\Box\varphi)$ since the way we use $\Vdash^+$ does not go into the term induction, but only formula induction within $Sb(\Box\varphi)$. Also, note that $\mathcal{E}(r^+, t)$ is identical with $\mathcal{E}(u, t)$ for any $u \in K$.

[38] Extend $\Vdash$ so that the new forcing relation preserves $\Vdash$'s feature of commuting with Booleans in the new state r^+. The old $\Vdash$ commutes with Booleans in all states in K.

On this model $\mathcal{K}^+$, we claim the following.

Claim 4.2.32 $\forall \psi \in Sb(\varphi)\ \forall u \in K^+\ [(u = r^+ \Longrightarrow u \Vdash^+ \psi$ *iff* $r \Vdash \psi)$ *and* $(u \neq r^+ \Longrightarrow u \Vdash^+ \psi$ *iff* $u \Vdash \psi)]$.

Proof Induction on the structure of ψ.

Case 1. ψ is atomic ($\psi = p$ or $\bot$). Immediate by definition.

Case 2. $\psi = \psi_1 * \psi_2$ ($* = \wedge, \vee$, or $\rightarrow$) or $\neg\psi_1$ (Boolean combinations) These cases are straightforward.

Case 3. $\psi = \Box\rho$.

Case 3.1. We first show the first conjunct of the statement of the claim (with a free variable for $u \in K^+$). Suppose $u = r^+$. We show $u \nVdash^+ \Box\rho$ iff $r \nVdash \Box\rho$.

$\Longleftarrow$) is straightforward. Suppose $\exists v \in K(rRv$ and $v \nVdash \rho)$. Call it v_1. However, $K \subseteq K^+$ and $R \subseteq R^+$. Also, since $u = r^+$, $v_1 \in K$ and $r^+ \notin K$, $v_1 \neq r^+$. By IH (using the second part with v_1), $v_1 \nVdash^+ \rho$ iff $v_1 \nVdash \rho$. So, $v_1 \nVdash^+ \rho$. Since r^+R^+r, these imply $\exists v \in K^+(uR^+v$ and $v \nVdash^+ \rho)$.

$\Longrightarrow$) Suppose $\exists v \in K^+(uR^+v$ and $v \nVdash^+ \rho)$. Call it v_1.

uR^+v_1 implies $u \neq v_1$. So, $r^+ \neq v_1$. So, $v_1 \in K$. Also, by IH and $v_1 \neq r^+$, $v_1 \nVdash^+ \rho$ iff $v_1 \nVdash \rho$. Hence, $v_1 \nVdash \rho$. Here we have two subcases.

Subcase 1. $v_1 = r$. This implies $r \nVdash \rho$. However, since the model $\mathcal{K}$ satisfies φ-soundness. So, $r \Vdash \Box\rho \rightarrow \rho$. This implies $r \nVdash \Box\rho$, as desired.

Subcase 2. $v_1 \neq r$. However, since $v_1 \in K$ and r is the root of $\mathcal{K}$, rRv_1. We have already shown $v_1 \nVdash \rho$. Hence, $\exists v \in K$ $(rRv$ and $v \nVdash \rho)$. So, $r \nVdash \Box\rho$.

Case 3.2. Now we move on to the second conjunct of the statement of the claim. Suppose $u \neq r^+$. We show $u \nVdash^+ \Box\rho$ iff $u \nVdash \Box\rho$.

$\Longleftarrow$) is straightforward. Suppose $\exists v \in K(uRv$ and $v \nVdash \rho)$. Call it v_1.

$v_1 \in K$ implies $v_1 \neq r^+$. By IH (the second part with v_1), $v_1 \nVdash^+ \rho$ iff $v_1 \nVdash \rho$. So, $v_1 \nVdash^+ \rho$. Since $K \subseteq K^+$ and $R \subseteq R^+$, $\exists v \in K^+(uR^+v$ and $v \nVdash^+ \rho)$ follows.

$\Longrightarrow$) Note first that $u \neq r^+$ implies $u \in K$. Suppose $\exists v \in K^+(uR^+v$ and $v \nVdash^+ \rho)$. Call it v_1. Since $u \in K$, uR^+v_1 implies $v_1 \neq r^+$. Hence, $v_1 \in K$. Thus, $(u, v_1) \in K \times K$. So, $(u, v_1) \in R^+ \cap (K \times K)$. So, $(u, v_1) \in R$. By IH and $v_1 \neq r^+$ implies $v_1 \Vdash^+ \rho$ iff $v_1 \Vdash \rho$. Hence, by assumption, $v_1 \nVdash \rho$ follows. Therefore, $\exists v \in K$ $(uRv$ and $v \nVdash \rho)$.

Case 4. $\psi = t : \rho$.

Case 4.1. We prove the first part. Suppose $u = r^+$ to show $u \nVdash^+ t : \rho$ iff $r \nVdash t : \rho$.

$\Longleftarrow$) Suppose $\rho \notin \mathcal{E}(r, t)$ or $\exists v \in K(rR^e v$ and $v \nVdash \rho)$. Call it v_1.

By definition of $\mathcal{E}^+$, the left disjunction implies $\rho \notin \mathcal{E}^+(r^+, t)$.

On the other hand, $K \subseteq K^+$ and $R^e \subseteq R^{e^+}$. Since $v_1 \in K$, $v_1 \neq r^+$. $rR^e v_1$ implies $rR^{e^+}v_1$. Also, $r^+R^{e^+}r$. By transitivity, $r^+R^{e^+}v_1$. Applying IH, $v_1 \nVdash^+ \rho$ iff $v_1 \nVdash \rho$. So, $v_1 \nVdash^+ \rho$. Putting them together, we get $\exists v \in K^+(uR^{e^+}v$ and $v \nVdash^+ \rho)$. Either way, $\rho \notin \mathcal{E}^+(r^+, t)$ or $\exists v \in K^+(uR^{e^+}v$ and $v \nVdash^+ \rho)$.

$\Longrightarrow$) Suppose $\rho \notin \mathcal{E}^+(r^+, t)$ or $\exists v \in K^+(uR^{e^+}v$ and $v \nVdash^+ \rho)$. The left disjunct derives $\rho \notin \mathcal{E}(r, t)$ by the definition of $\mathcal{E}^+$.

For the right disjunct, we start from $\exists v \in K^+(uR^{e^+}v$ and $v \nVdash^+ \rho)$. Call it v_1. $v_1 \in K^+$, $uR^{e^+}v_1$, and $v_1 \nVdash^+ \rho$. Here we have two subcases.

Subcase 1. $v_1 \in K$. $v_1 \in K$ implies $v_1 \neq r^+$. So, by IH, $v_1 \nVdash^+ \rho$ iff $v_1 \nVdash \rho$. So, $v_1 \nVdash \rho$. Also, $v_1 \in K$ implies rRv_1 or $v_1 = r$ (since r is the root of $\mathcal{K}$). But either way, rR^ev_1 follows (in the first case, $R \subseteq R^e$ and in the second case $v_1 = r$ and $v_1R^ev_1$ implies it). These imply $\exists v \in K$ $(rR^ev$ and $v \nVdash \rho)$.

Subcase 2. $v_1 \in K^+ \backslash K$. This implies $v_1 = r^+$. By IH (using the first part), $v_1 \nVdash^+ \rho$ iff $r \nVdash \rho$. Hence, $r \nVdash \rho$ follows. By reflexivity of R^e, rR^er. Since $r \in K$, $\exists v \in K$ $(rR^ev$ and $v \nVdash \rho)$.

Case 4.2. Now we move on to the second conjunct of the entire statement of the claim. Suppose $u \neq r^+$. We show $u \nVdash^+ t : \rho$ iff $u \nVdash t : \rho$.

$\Longleftarrow$) Suppose $\rho \notin \mathcal{E}(r,t)$ or $\exists v \in K(uR^ev$ and $v \nVdash \rho)$.

The first disjunct implies $\rho \notin \mathcal{E}^+(r^+,t)$ by the definition of $\mathcal{E}^+$.

For the second disjunct, call the existing state v_1. $v_1 \in K$ implies $v_1 \neq r^+$. Applying IH, we get $v_1 \nVdash^+ \rho$ iff $v_1 \nVdash \rho$. So, $v_1 \nVdash^+ \rho$. Together with $K \subseteq K^+$ and $R^e \subseteq R^{e^+}$, these imply that $\exists v \in K^+(uR^{e^+}v$ and $v \nVdash^+ \rho)$. Either way, $\rho \notin \mathcal{E}^+(r^+,t)$ or $\exists v \in K^+(uR^{e^+}v$ and $v \nVdash^+ \rho)$.

$\Longrightarrow$) Suppose $\rho \notin \mathcal{E}^+(r^+,t)$ or $\exists v \in K^+(uR^{e^+}v$ and $v \nVdash^+ \rho)$. The left disjunct derives $\rho \notin \mathcal{E}(r,t)$ by the definition of $\mathcal{E}^+$.

On the other hand, we start from $\exists v \in K^+(uR^{e^+}v$ and $v \nVdash^+ \rho)$. Call it v_1. $v_1 \in K^+$, $uR^{e^+}v_1$, and $v_1 \nVdash^+ \rho$. Here we have two subcases.

Subcase 1. $v_1 \in K$. This immediately implies $v_1 \neq r^+$. Hence, by IH, $v_1 \nVdash^+ \rho$ iff $v_1 \nVdash \rho$. So, $v_1 \nVdash \rho$. Also, by the assumption of this case, $u \neq r^+$.

This also implies $u \in K$. So, $u \in K$ and $v_1 \in K$. So $(u, v_1) \in K \times K$, So $(u, v_1) \in R^{e^+} \cap (K \times K)$. Hence, $uR^e v_1$. Combining these, we can derive $\exists v \in K(uR^e v$ and $v \nVdash \rho)$.

Subcase 2. $v_1 \in K^+ \backslash K$. This immediately implies $v_1 = r^+$. Applying IH, $v_1 \nVdash^+ \rho$ iff $r \nVdash \rho$. So, $r \nVdash \rho$. Also, $uR^{e^+} v_1$ implies $uR^{e^+} r^+$. However, since $r^+ R^+ r$ and $R^+ \subseteq R^{e^+}$, $r^+ R^{e^+} r$. So, by transitivity, $uR^{e^+} r$.

Note that $u \neq r^+$ under the current assumption. So, $u \in K$. Since obviously $r \in K$, $(u, r) \in K \times K$. Hence $(u, r) \in R^{e^+} \cap (K \times K)$. So, $uR^e r$. Therefore, $\exists v \in K(uR^e v$ and $v \nVdash \rho)$. ⊠ (claim)

This claim gives the following observation. Since $\mathcal{N}(1) = r$ and by the assumption for the proof of the whole lemma, i.e. $\mathcal{N}(1) \nVdash \varphi$, we have $r \nVdash \varphi$. By the claim (the second part), $r \nVdash^+ \varphi$. Since $r^+ R^+ r$ holds, $\mathcal{K}^+, r^+ \nVdash \Box\varphi$.

Now we check that our model $\mathcal{K}^+$ satisfies $\Box\varphi$-soundness condition, and also we extend the interpretation function appropriately.

Claim 4.2.33 *$\mathcal{K}^+$ is $\Box\varphi$-sound, i.e., $r^+ \Vdash^+ \Box\psi \to \psi$ for all $\Box\psi \in Sb(\Box\varphi)$.*

Proof Since $\mathcal{K}$ is φ-sound, $r \Vdash \Box\psi \to \psi$ for all $\Box\psi \in Sb(\varphi)$. Also, since $r \Vdash \Box\psi \to \psi$ iff $r \nVdash \Box\psi$ or $r \Vdash \psi$. The above claim implies $r^+ \nVdash^+ \Box\psi$ or $r^+ \Vdash^+ \psi$ for all $\Box\psi \in Sb(\varphi)$, so $r^+ \Vdash^+ \Box\psi \to \psi$ for all $\Box\psi \in Sb(\varphi)$. By the above argument, $\mathcal{K}, r \nVdash \varphi$ implies $\mathcal{K}^+, r^+ \nVdash^+ \Box\varphi$. But this implies $r^+ \Vdash^+ \Box\varphi \to \varphi$. Hence, we have $r^+ \Vdash \Box\psi \to \psi$ for all $\Box\psi \in Sb(\Box\varphi)$. ⊠ (claim)

Let us redefine a new (extended) interpretation function $\mathcal{N}^+$ over the extended model.[39] Let $\mathcal{N}^+(1) = r^+$ and $\mathcal{N}^+(1.\sigma) = \mathcal{N}(\sigma)$. So, $\mathcal{K}^+, \mathcal{N}^+(1) \nVdash^+ \Box\varphi$. Hence, $F\Box\varphi$ is realized in $\mathcal{N}^+(1)$ in $\mathcal{K}^+$. Thus, $\{1F\Box\varphi\}$ is satisfiable. Also, by the last claim, $\mathcal{K}^+$ is $\Box\varphi$-sound. ⊠(Lemma)

Soundness of TGLA can be shown as follows. Note that the proof of soundness is different from usual cases of normal modal logics. Our models are φ-sound models and we cannot freely take a generated submodel of a given model since there is no guarantee that the resulting model is also a φ-sound model.

Theorem 4.2.34 (Soundness for TGLA) *If φ has a prefixed GLA-tableau proof, then φ is valid in all φ-sound GLA models.*

Proof We first show the soundness for $\Box\varphi$-sound models, and then show the soundness for φ-soundness.

Suppose, for the sake of reductio, φ has a tableau proof in TGLA, namely, $\{1F\Box\varphi\}$ has a closed tableau in TGLA and that $\Box\varphi$ is invalid. Then, there exists a $\Box\varphi$-sound GLA-model, s.t. $\exists u \in K$, s.t. $u \nVdash \Box\varphi$.

For u, we have $u = \mathcal{N}(1)$ or $u \neq \mathcal{N}(1)$. In the former case, $\mathcal{N}(1) \nVdash \Box\varphi$. So, $\{1F\Box\varphi\}$ is satisfiable (in a $\Box\varphi$-sound GLA-model). Since all the rules

---

[39] We modify $\mathcal{N}$ so that $1 < \sigma \Longrightarrow \mathcal{N}^+(1)R^+\mathcal{N}^+(\sigma)$ (for $\sigma \neq 1$). This interpretation has to satisfy the properties of GLA-interpretation. $\mathcal{N}^+(1) = r^+$, so irreflexivity is preserved in extension since r^+ is different from any $u \in K$. Let σ_1 and σ_2 be such that $\mathcal{N}^+(1)R^+\mathcal{N}^+(\sigma_1)$ and $\mathcal{N}^+(\sigma_1)R^+\mathcal{N}^+(\sigma_2)$. Then by transitivity of R^+, we have $\mathcal{N}^+(1)R^+\mathcal{N}^+(\sigma_2)$. So, the interpretation function $\mathcal{N}^+$ is well-defined for R^+. For R^{e^+}, the reflexivity and symmetry cases are obvious, and the case of transitivity is similar to R^+.

of TGLA preserve satisfiability[40], any tableau constructed from $\{1F\Box\varphi\}$ is satisfiable. But no satisfiable tableau is closed. Contradiction.

In the latter case, since $\mathcal{N}(1)$ is the root node, $\mathcal{N}(1)Ru$. On the other hand, $u \nVdash \Box\varphi$ implies that there exists $v \in K$, s.t. uRv and $v \nVdash \varphi$.

By transitivity, $\mathcal{N}(1)Ru$ and uRv implies $\mathcal{N}(1)Rv$. $\mathcal{N}(1)Rv$ and $v \nVdash \varphi$ implies $\mathcal{N}(1) \nVdash \Box\varphi$. So, $\{1F\Box\varphi\}$ is satisfiable (in a $\Box\varphi$-sound GLA-model). Since all the rules of TGLA preserve satisfiability, any tableau constructed from $\{1F\Box\varphi\}$ is satisfiable. But no satisfiable tableau is closed. Contradiction.

Hence, $\{1F\Box\varphi\}$ has a closed tableau TGLA $\Longrightarrow$ $\Box\varphi$ is valid in all $\Box\varphi$-sound GLA models.

However, by the lemma 4.2.31, $\Box\varphi$ is valid in all $\Box\varphi$-sound GLA models $\Longrightarrow$ φ is valid in all φ-sound GLA models. Therefore, $\{1F\Box\varphi\}$ has a closed tableau TGLA $\Longrightarrow$ φ is valid in all φ-sound GLA models. ⊠

Next, we show weak completeness of TGLA for a formula φ with respect to φ-sound Kripke models. We prove completeness of TGLA by taking the following steps.

(1) Assume unprovability of φ, i.e. $\{1F\Box\varphi\}$ has no closed tableau in TGLA.

(2) We prove that this implies $\{1F\Box^{N+1}\varphi\}$ has no closed tableau in TwGLA. (Here N stands for $|\{\Box\psi | \Box\psi \in Sb(\varphi)\}|$.)

---

[40] In the case of Reflection Rule, we shift from one model to another, but this does not raise any serious difference from other rules since what matters here is satisfiability (which is stated in an existential form).

(3) By using completeness of TwGLA, we can show that there exists a GLA-model $\mathcal{K}$ such that $\mathcal{K}, r \nVdash \Box^{N+1}\varphi$.

(4) By using the argument that has been used in the literature of provability logic (originally going back to Solovay and Guaspari [78]), we transform the model constructed in (3) into a φ-sound GLA-model.

This is an outline of our proof of completeness of TGLA. Here (2) $\Longrightarrow$ (3) is just completeness of TwGLA. We show the details of our proof of completeness in the following. At first, we show (1) $\Longrightarrow$ (2) as follows.

Proposition 4.2.35 *If there is a closed tableau for* $1F\Box^n\varphi$ $(n > 1)$ *in TwGLA, then* $1F\Box\varphi$ *has a closed tableau in TGLA.*

Proof Suppose $1F\Box^n\varphi$ has a closed tableau in TwGLA $(n > 1)$.

Pick one. Then apply reflection rule $n-1$ times on top of the closed tableau. The resulting tableau must be a closed tableau for $1F\Box\varphi$ in TGLA. $\boxtimes$

Now we want to prove the weak completeness of GLA. (This involves steps from (2) to (4). We indicate which part of the proof corresponds to which step in the following.) Our proof is based on a proof given in [4].

Theorem 4.2.36 (Completeness for GLA) *If* φ *has no tableau proof in TGLA, i.e.,* $\{1F\Box\varphi\}$ *has no closed tableau in TGLA, then there exists a* φ*-sound GLA Kripke model* $\mathcal{K}$ $(= (K, R, R^e, r, \mathcal{E}, \Vdash))$ *and there exists* $u \in K$*, s.t.* $\mathcal{K}, u \nVdash \varphi$*.*

Proof Suppose $1F\Box\varphi$ has no closed tableau in the prefixed tableau system for GLA. In particular, this implies that there is also no closed tableau for this in the prefixed tableau system for wGLA. By completeness of wGLA, we do have a wGLA-countermodel for φ. However, the problem is that this does not guarantee that the model is a φ-sound GLA model that falsifies φ. The following argument is given in order to guarantee this condition.

Let $N = |\{\Box\psi | \Box\psi \in Sb(\varphi)\}|$. Proposition 4.2.35 (i.e., (1) $\Rightarrow$ (2)). implies that if $1F\Box^{N+1}\varphi$ has a closed tableau in wGLA, then $1F\Box\varphi$ has a closed tableau in GLA. So, taking the contrapositive, by the assumption, we obtain the statement $1F\Box^{N+1}\varphi$ has no closed tableau in wGLA. By completeness of wGLA, this implies that there exists a (not necessarily φ-sound) GLA-model $\mathcal{K}$ s.t. $\mathcal{K}, r \nVdash \Box^{N+1}\varphi$. (This corresponds to the step (2) $\Rightarrow$ (3) in the outline.)

Now we show the step (3) $\Rightarrow$ (4) in the outline. $\mathcal{K}, r \nVdash \Box^{N+1}\varphi$ implies that there is a sequence of nodes $r = a_0Ra_1Ra_2 \dots a_NRa_{N+1}$, s.t. $a_i \nVdash \Box^{N+1-i}\varphi$ where $0 \leq i \leq N+1$. (Note that $a_{N+1} \nVdash \varphi$.)

None of the formulas $\Box\psi \to \psi$ ($\Box\psi \in Sb(\varphi)$) can be false at two (or more) different nodes a_i and a_j. Indeed, suppose there is such a ψ_1. Then, $a_i \nVdash \Box\psi_1 \to \psi_1$ and $a_j \nVdash \Box\psi_1 \to \psi_1$, where we assume $a_i < a_j$ without loss of generality. Then, $a_i \Vdash \Box\psi_1$ and $a_i \nVdash \psi_1$. Also, $a_j \Vdash \Box\psi_1$ and $a_j \nVdash \psi_1$. But this clearly raises a contradiction.

We have N formulas of the form $\Box\psi \to \psi$, each of which is false at most one world in a chain of $N+1$-many world. Then, there exists at least one world a_i in which $\Box\psi \to \psi$ are true at a_i. Namely, $\exists i$ ($0 \leq i \leq N+1$), s.t.

$a_i \Vdash \bigwedge\{\Box\psi \to \psi | \Box\psi \in Sb(\varphi)\}$. We may have either a case $a_i R a_{N+1}$ or a case $a_i = a_{N+1}$. Either way, we take the restriction of $\mathcal{K}$. More precisely, take such a subset $K'(= \{u \in K | a_i R u \text{ or } a_i = u\})$ as the underlying set of the Kripke model that we want to construct, take the two accessibility relation as the restrictions of R and R^e by this set accordingly, take the evidence function as the restriction of $\mathcal{E}$ concerning its first coordinate to the above set, take the restriction of forcing relation, and let the new root $r' = a_i$. We call the new model $\mathcal{K}' = (K', R', R^{e\prime}, r', \mathcal{E}', \Vdash')$. Then, by construction, $\mathcal{K}'$ is a φ-sound GLA model. Note that this new root was already a part of a model of wGLA, s.t. $r' \nVdash \Box^{N+1-i}\varphi$. Also, this truncation does not affect the structure of that part of the model which is used to ensure that for some $u \in K'$, $u \nVdash \varphi$ since the restriction is made below a_{N+1} and the structure of the the accessibility relation R in the canonical model (above a_{N+1}) is determined only by the structure of the formula φ.[41] Except the case of $R^{e\prime}$, all other conditions of a model of GLA are clearly satisfied. For $R^{e\prime}$, this is the restriction of R^e to $K' \times K'$, i.e., $R^{e\prime} = R^e \cap (K' \times K')$. It is easy to check that this is still an equivalence relation and $R' \subseteq R^{e\prime}$. So, $R^{e\prime}$ is also taken care of. Therefore, the restriction $\mathcal{K}'$ is a φ-sound GLA model such that for some $v \in K'$, $v \nVdash \varphi$. ⊠

---

[41] At this point, the $\Box$ added to deal with soundness is not important (and can be taken as a part of $\Box^{N+1}$), so we directly discuss a state u, s.t. $u \nVdash \varphi$ (instead of $\Box\varphi$).

4.2.3 Cut-admissibility of Hypersequent Calculi for **GrzA** and **GLA**

We eventually translate proofs in our prefixed tableau systems to proofs in hypersequent calculi for GrzA and GLA. We first present systems and then show how we translate prefixed tableau systems to hypersequent calculi.

4.2.3.1 Hypersequent Calculi for **GrzA** and **GLA**

In our formulation in this paper, a hypersequent is a finite set of sequents in a traditional Gentzen-style sequent calculus, which is written as follows. $\Gamma_1 \Rightarrow \Delta_1 | \dots | \Gamma_n \Rightarrow \Delta_n$, where "|" has a disjunctive reading. (The precise definition of this will be given later.) To make our translation from the prefixed tableau system smooth, we take a sequent to be sets of formulas. So, a hypersequent is a set of sets of formulas. Because of this formulation, some structural rules, i.e., internal and external exchange rules, internal and external contraction rules, can be omitted from our systems.

Due to the presence of Reflection Rule in GLA, we first formulate a hypersequent calculus for GrzA (HGrzA) and one for a system wGLA (HwGLA), and then we present the hypersequent calculus that exactly corresponds to GLA (HGLA). The weaker systems look as follows.

I. Common rules for HGrzA and HwGLA.[42]

---

[42] Since the languages in the modal logics are different, these rules are, strictly speaking, not common. However, we assume that these rules are stated *schematically*. We consider

1) **Axiom**: $A \Rightarrow A$ $\qquad$ $\bot \Rightarrow$

2) **External structural rules**: $\mathbf{EW}\ \dfrac{G}{G|H}$.

3) **Internal structural rules**:

$$\mathbf{LW}\ \frac{G|\Gamma \Rightarrow \Delta}{G|A, \Gamma \Rightarrow \Delta} \qquad \mathbf{RW}\ \frac{G|\Gamma \Rightarrow \Delta}{G|\Gamma \Rightarrow \Delta, A}$$

4)**Operational rules**

$$\mathbf{L}\wedge\ \frac{G|A, B, \Gamma \Rightarrow \Delta}{G|A \wedge B, \Gamma \Rightarrow \Delta} \qquad \mathbf{R}\wedge\ \frac{G|\Gamma \Rightarrow \Delta, A \qquad G|\Gamma \Rightarrow \Delta, B}{G|\Gamma \Rightarrow \Delta, A \wedge B}$$

$$\mathbf{L}\vee\ \frac{G|A, \Gamma \Rightarrow \Delta \quad G|B, \Gamma \Rightarrow \Delta}{G|A \vee B, \Gamma \Rightarrow \Delta} \qquad \mathbf{R}\vee\ \frac{G|\Gamma \Rightarrow \Delta, A, B}{G|\Gamma \Rightarrow \Delta, A \vee B}$$

$$\mathbf{L} \rightarrow\ \frac{G|\Gamma \Rightarrow \Delta, A \qquad G|B, \Gamma \Rightarrow \Delta}{G|A \rightarrow B, \Gamma \Rightarrow \Delta} \qquad \mathbf{R} \rightarrow\ \frac{G|A, \Gamma \Rightarrow \Delta, B}{G|\Gamma \Rightarrow \Delta, A \rightarrow B}$$

---

all substitution instances of the rules restricted to each language.

L¬ $\dfrac{G|\Gamma \Rightarrow \Delta, A}{G|\Gamma, \neg A \Rightarrow \Delta}$ **R¬** $\dfrac{G|\Gamma, A \Rightarrow \Delta}{G|\Gamma \Rightarrow \Delta, \neg A}$

5) **Rules for Proof-terms of LP**

Lt $\dfrac{G|A, \Gamma \Rightarrow \Delta}{G|t : A, \Gamma \Rightarrow \Delta}$ **R+** $\dfrac{G|\Gamma \Rightarrow \Delta, t : A}{G|\Gamma \Rightarrow \Delta, t + s : A}$ $\dfrac{G|\Gamma \Rightarrow \Delta, s : A}{G|\Gamma \Rightarrow \Delta, t + s : A}$

R! $\dfrac{G|\Gamma \Rightarrow \Delta, t : A}{\Gamma \Rightarrow \Delta, !t : t : A}$ **R·** $\dfrac{G|\Gamma \Rightarrow \Delta, t : A \to B \qquad G|\Gamma \Rightarrow \Delta, s : A}{G|\Gamma \Rightarrow \Delta, t \cdot s : B}$

Labeled Splitting $\dfrac{G|\vec{t} : \Gamma_1, \Gamma_2 \Rightarrow \vec{s} : \Delta_1, \Delta_2}{G|\vec{t} : \Gamma_1 \Rightarrow \vec{s} : \Delta_1 | \Gamma_2 \Rightarrow \Delta_2}$

6) **Cut** $\dfrac{G|\Gamma \Rightarrow \Delta, A \qquad G|A, \Gamma \Rightarrow \Delta}{G|\Gamma \Rightarrow \Delta}$

II. **Rule for GL** $\Box$ $\quad \Box GL$ $\dfrac{G|\Box\varphi, \Box\Gamma, \Gamma \Rightarrow \varphi}{G|\Box\Gamma \Rightarrow \Box\varphi}$

III. **Rules for Grz** $\boxdot$ $\quad L\boxdot$ $\dfrac{G|A, \Gamma \Rightarrow \Delta}{G| \boxdot A, \Gamma \Rightarrow \Delta}$

$$\boxdot Grz \; \frac{G| \boxdot (A \to \boxdot A), \boxdot \Gamma \Rightarrow A}{G| \boxdot \Gamma \Rightarrow \boxdot A}$$

IV. **Constant Specification Rule for wGLA** $\quad \dfrac{}{\Rightarrow c : A}$

where A is an axiom of wGLA and $A \in CS(c)$ (CS is for wGLA).

V. **Constant Specification Rule for GrzA** $\quad \dfrac{}{\Rightarrow c : A}$

where A is an axiom of GrzA and $A \in CS(c)$ (CS is for GrzA).

1. The hypersequent calculus consisting of rules I (formulated in the language of GLA) + II + IV is the hypersequent calculus for wGLA, i.e. HwGLA.

2. The hypersequent calculus consisting of rules I (formulated in the language of GLA) + III + V is the hypersequent calculus for GrzA, i.e. HGrzA.

By adding the following rule to HwGLA and changing the constant specification rule accordingly, we obtain a system that exactly corresponds to GLA, i.e. HGLA.

Reflection Rule (RR) $\quad \dfrac{\Rightarrow \Box\varphi}{\Rightarrow \varphi}$

Remark: By using the Reflection Rule, we can derive $t : \Box\varphi \to \varphi$ as follows. Note that the $\Box(t : \Box\varphi \to \varphi)$ can be proven without using the Reflection Rule.

$$
\begin{array}{c}
\varphi \Rightarrow \varphi \\
\hline\hline
\Rightarrow t : \Box\varphi, \Box(t : \Box\varphi \to \varphi), \Box\varphi, \varphi \Rightarrow \varphi \\
\hline
\Rightarrow \Box(t : \Box\varphi \to \varphi), \Box\varphi, \varphi \Rightarrow t : \Box\varphi \to \varphi \\
\hline
\Rightarrow \Box\varphi \Rightarrow \Box(t : \Box\varphi \to \varphi) \\
\hline
\Rightarrow t : \Box\varphi \Rightarrow \Box(t : \Box\varphi \to \varphi) \\
\hline
\Rightarrow t : \Box\varphi \Rightarrow \; | \Rightarrow \Box(t : \Box\varphi \to \varphi) \\
\hline
t : \Box\varphi \Rightarrow \; | \Rightarrow \Box(t : \Box\varphi \to \varphi) \\
\hline
t : \Box\varphi \Rightarrow \varphi | \Rightarrow \Box(t : \Box\varphi \to \varphi) \\
\hline
\Rightarrow t : \Box\varphi \to \varphi | \Rightarrow \Box(t : \Box\varphi \to \varphi) \\
\hline
\Box(t : \Box\varphi \to \varphi) \Rightarrow t : \Box\varphi \to \varphi | \Rightarrow \Box(t : \Box\varphi \to \varphi) \\
\hline
\Rightarrow \Box(t : \Box\varphi \to \varphi) | \Rightarrow \Box(t : \Box\varphi \to \varphi) \\
\hline
\dfrac{\Rightarrow \Box(t : \Box\varphi \to \varphi)}{\Rightarrow t : \Box\varphi \to \varphi} \text{Reflection Rule}
\end{array}
$$

Theorem 4.2.37 (Deductive Equivalence) *HGLA*$\vdash \Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_n \Rightarrow \Delta_n$ *if and only if GLA*$\vdash \Box(\bigwedge \Gamma_1 \to \bigvee \Delta_1) \vee \cdots \vee \Box(\bigwedge \Gamma_n \to \bigvee \Delta_n)$.

Proof In both ways, a proof is given by the length of derivation in two systems, respectively. $\Longleftarrow$ is straightforward. $\Longrightarrow$ needs some discussions. We highlight a few important cases. One is Labeled Splitting. To derive Labeled Splitting in the Hilbert-style system, we can use the fact GLA$\vdash (t_1 : A \to t_2 : B) \to \Box(t_1 : A \to t_2 : B)$. Another one is the identification of HGLA$\vdash \Gamma_1 \Rightarrow \Delta_1 | \Gamma_1 \Rightarrow \Delta_1$ with HGLA$\vdash \Gamma_1 \Rightarrow \Delta_1$ due to the set-hood of hypersequents.[43] $\mathsf{HGLA}\vdash \Gamma_1 \Rightarrow \Delta_1 | \Gamma_1 \Rightarrow \Delta_1$ is translated into $\mathsf{GLA}\vdash \Box(\bigwedge \Gamma_1 \to \bigvee \Delta_1) \vee \Box(\bigwedge \Gamma_1 \to \bigvee \Delta_1)$. Hence, by propositional logic $\mathsf{GLA}\vdash \Box(\bigwedge \Gamma_1 \to \bigvee \Delta_1)$. Since the Hilbert-style axiom system for GLA has Reflection Rule, we get $\mathsf{GLA}\vdash \bigwedge \Gamma_1 \to \bigvee \Delta_1$.

For the case of Reflection Rule in HGLA, we can argue as follows. Since these do not have any side sequent for Reflection Rule and we do not have to put $\Box$ for $|$, $\Rightarrow \Box\varphi$ and $\Rightarrow \varphi$ can be translated into $\Box\varphi$ and φ, so derivability of the translated rule in the Hilbert-style is obvious. ⊠

We can prove an analogous theorem on deductive equivalence for HGrzA. The proof is similar to that of HGLA and even simpler.

Theorem 4.2.38 *$\mathsf{HGrzA}\vdash \Gamma_1 \Rightarrow \Delta_1 | \dots | \Gamma_n \Rightarrow \Delta_n$ if and only if $\mathsf{GrzA}\vdash \boxdot(\bigwedge \Gamma_1 \to \bigvee \Delta_1) \vee \cdots \vee \boxdot(\bigwedge \Gamma_n \to \bigvee \Delta_n)$.*

4.2.3.2 Translation from Prefixed Tableaux to Hypersequents

4.2.3.2.1 Definition of Translation

The deductive equivalence theorem presented in the previous subsection uses cut, so it does not tell us whether this hypersequent calculus enjoys cut-admissibility or not. So, we next show that the systems HGLA and HGrzA are indeed complete without using cut. Combined with soundness of HGLA and HGrzA with cut, which follows from the equivalence of them to Hilbert-style

---

[43] The problematic direction is the one corresponding to external contraction, which is absorbed in the set-hood of hypersequents. The other direction is obvious.

systems for GLA and GrzA, this suffices to prove cut-admissibility semantically. The proof goes by constructing a translation of prefixed tableau proofs to proofs in the hypersequent calculi.

First let us mention how to handle Reflection Rule in HGLA. If we used a usual format of prefixed tableau system, then we would have a situation in which although we have no syntactic difference between the two prefixes, semantically they are interpreted differently (these are interpreted as the roots of the two different models, rather than two states in the same model). This would raise a slight complication to our translation since the translation relies on the syntactic feature of the prefixed tableau system. To avoid such a situation, we use a format that looks like a destructive tableau rule for Reflection Rule. This makes simpler our translation from the prefixed tableau system to the hypersequent. Once we turn "upside down"[44] the reflection rule in the prefixed tableau system, i.e. from $\{1F\Box\varphi\}$, infer $\{1F\varphi\}$, we can naturally transform this rule into the following rule in our hypersequent calculus. (Note that we do not have any side sequents in this rule.)

$$\frac{\Rightarrow \Box\varphi}{\Rightarrow \varphi}$$

Next, we define a translation from the deductive meta-language[45], of the

---

[44] This may not look like simply upside down any more. We can call this "reversing the order of application".

[45] This terminology is due to Kosta Došen [42] (see Chapter 2). We call "deductive-metalanguage" a (semi-)formal language that is used to formulate a deductive system such as a sequent calculus or a tableau system. We do not use induction on symbols in a

prefixed tableau system to that of the hypersequent calculus. By using that, we show that the derivability in the prefixed tableau system is preserved under the translation in the hypersequent calculi. Our translation mapping, which is called s, is defined in two stages.

First, we define a mapping t that maps a set of prefixed formulas to a set of sets of signed formulas in the following way. (Although the definition is essentially the same as the case of S4LPN, we give the definition again since this is important.)

1. The set of prefixed formulas is partitioned into subsets so that all formulas with the same prefixes σ_i go into the same subset.

2. We strip off prefixes from those partitioned prefixed formulas (for each σ_i).

3. We call the resulting set H_{σ_i} for each σ_i, i.e. $H_{\sigma_i} := \{\Phi | \sigma_i \Phi \in S\}$.

4. We arrange those H_{σ_i} by using some order[46] via "|" in the hypersequent calculus. So, we have $H_1 | \dots | H_{\sigma_1} | H_{\sigma_2} | \dots | H_{\sigma_n}$. Our reading "|" is the same as that of hypersequent, so we have now constructed a set of sets of signed formulas.[47] Let $S^t := H_1 | \dots | H_{\sigma_1} | H_{\sigma_2} | \dots | H_{\sigma_n}$.

Second, we consider a mapping that maps a set of sets of signed formulas, i.e. $H_1 | \dots \ | H_{\sigma_1} | H_{\sigma_2} | \dots | H_{\sigma_n}$, to a set of sequents. This map-

---

formula in the language, but the language is different from the background meta-language that is used to formulate and prove theorems here.

[46] The order can be arbitrary since our hypersequents are sets of sets of formulas.

[47] Each H_{σ_i} will work as each sequent occurring in the hypersequent that we will obtain as an image of the mapping s we are defining. But note that each H_{σ_i} consists of only signed formulas.

ping can be readily constructed by putting T formulas to the antecedent and F formulas to the succedent for each case of H_{σ_i}. Namely, if $H_{\sigma_i} = \{T\varphi_1, \dots, T\varphi_k, F\psi_1, \dots, F\psi_m\}$, then we map this to $\varphi_1, \dots, \varphi_k \Rightarrow \psi_1, \dots, \psi_m$. This is the only thing that we have to do for each H_{σ_i}, but, for notational simplicity, we officially define mapping for an entire set of H_{σ_i}'s. We call this mapping u.

$(H_1|\dots|H_{\sigma_1}|H_{\sigma_2}|\dots|H_{\sigma_n})^u = \varphi_{1,1}, \dots, \varphi_{1,k_{\sigma_1}} \Rightarrow \psi_{1,1}, \dots, \psi_{1,m_{\sigma_1}}|\varphi_{\sigma_1,1}, \dots, \varphi_{\sigma_1,k_{\sigma_1}} \Rightarrow \psi_{\sigma_1,1}, \dots, \psi_{\sigma_1,m_{\sigma_1}}|\dots|\varphi_{\sigma_n,1}, \dots, \varphi_{\sigma_n,k_{\sigma_n}} \Rightarrow \psi_{\sigma_n,1}, \dots, \psi_{\sigma_n,m_{\sigma_n}}$.

We finally define the desired mapping s as a composition of t and u, i.e., $S^s := ((S^t)^u)$. So, S^s will be of the form $\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_{\sigma_i} \Rightarrow \Delta_{\sigma_i}|\dots|\Gamma_{\sigma_n} \Rightarrow \Delta_{\sigma_n}$.

4.2.3.2.2 Potential problems and an outline of their solution

We need some care in order to guarantee that the translation defined above from the prefixed tableau system for GLA to the hypersequent calculus for GLA. (The case of GrzA is similar.)

The diagnosis of the difficulty in translating any proof in the prefixed tableau system to a hypersequent proof is essentially the same as the case of S4LPN. The same sources of the problems we pointed out as i) and ii) for S4LPN matter here, and in ii) we can classify two different problems 1) and 2), as follows.

i) π-rule is applied earlier than α, β-rule in a prefixed tableau, so in the translated system, we miss some cases of possible applications of $\Box R$ corresponding to ν-rule with new prefixes (since the outermost logical symbols of the relevant formulas are not $\Box$.) (E.g. $\{1T\Box A \vee \Box B, 1F\Box(A \vee B)\}$.)

ii) After π-rule is applied, some formula(s) with proof terms move(s) back to previous worlds by reverse rules, and $\nu_{K,4}$ rules are applied subsequently.

Here is a counterexample against a naive translation in TGLA, due to case ii).

$1F\Box(t : \Box A \to A)$

$1.1.Ft : \Box A \to A$; $1.1T\Box(t : \Box A \to A)$

$1.1.Tt : \Box A$; $1.1.FA$

$1Tt : \Box A$; $1T\Box A$

$\mathbf{1.1TA}$; $\mathbf{1.1T\Box A}$

×

One step of the translation of this prefixed tableau proof is given as follows.

$$\frac{t : \Box A, \Box A \Rightarrow \Box(t : \Box A \to A) | \Box(t : \Box A \to A), t : \Box A, A \Rightarrow t : \Box A \to A, A}{t : \Box A, \Box A \Rightarrow \Box(t : \Box A \to A) | \Box(t : \Box A \to A), t : \Box A \Rightarrow t : \Box A \to A, A} ?$$

Since there is no rule in hypersequent calculi that exactly corresponds to

ν_K, there is no way of directly justifying this step in HGLA (this example is essentially the same as the one we give before except that π-rule here is π-rule for GL). We have two problems. 1) Reverse rules produce a context in which $\nu_{K,4}$ have to be applied separately from π. 2) Reverse rules make a prefix used more than once when we apply $\nu_{K,4}$ (see the example above).

However, it may still be that any prefixed tableau proof can be transformed into some prefixed tableau proof that can be translated into a hypersequent. Indeed, this is what we do in order to solve the problems. Like in the case of S4LPN, we solve the two problems by showing that any (cut-free) prefixed tableau proof can be turned into a π-$\nu_{K,4}$ normal form via the "proof confluence property" of prefixed tableau systems and preservation of closure under taking numerical variants of tableaux.

Then, to solve the problem that there is no rule in hypersequent calculi corresponding to $\nu_{K,4}$, again like in the case of S4LPN, we translate a proof in a π-$\nu_{K,4}$-normal form in TGLA into a variant of the system in which π-rule for GL is so modified that $\nu_{K,4}$-rules are not primitive but admissible. (The case of π-rule for GrzA is similar.) Here the π-rules are eventually modified in the following way.

$\pi^\sharp$-rule for GL : Let S be a set of prefixed formulas (a tableau constructed up to the previous step). Let $\sigma.nF\varphi$ be the conclusion of the traditional π-rule ($\sigma.n$ is new on a branch), and $S^\sharp_{\sigma.n} = \{\sigma.nT\Box\psi, \sigma.nT\psi | \sigma T\Box\psi \in S\}$. (For Grz, use the similar set with $\boxdot$.)

Then we have the following rule. $$\frac{\{\sigma F\Box\varphi\}\cup S}{\{\sigma.nF\varphi, \sigma.nT\Box\varphi\}\cup S^{\sharp}_{\sigma.n}}$$

π-rule for Grz is as follows (with $\boxdot$). $$\frac{\{\sigma F\boxdot\varphi\}\cup S}{\{\sigma.nF\varphi, \sigma.nT\boxdot(\varphi\to\boxdot\varphi)\}\cup S^{\sharp}_{\sigma.n}}$$

These rules combine the idea of constructing a new prefix in the prefixed tableau and the idea of destructive tableau. (The case of π rule for Grz is similar to π rule for GL, so we discuss only on GLA case.) The π rule for GL carries $\Box\psi, \psi$ to the new world automatically. (In an auxiliary proof system we eventually introduce, these rule may spoil "strong proof confluence." Hence, we treat this rule not as a primitive rule in TGLA, but we introduce a variant of TGLA in which $\pi^{\sharp}$ is treated as primitive so that we can use the property of proof confluence in the original system TGLA.)

Combining these ideas, we translate a proof TGLA into a proof in HGLA. Our immediate goal is to obtain a method of manipulating an arbitrary (cut-free) proof in HGLA into a proof satisfying the condition: all applications of $\nu_{K,4}$-rules are made immediately after an application of π-rule where the application of $\nu_{K,4}$ uses the prefix introduced by the π rule directly preceding it. And we move on to $\pi^{\sharp}$.

4.2.3.2.3. Proof confluence in prefixed tableau systems

The structure of the proof of the modification of tableau proofs is almost the same as the case of S4LPN. Thus, we write again important definitions but omit some details of the proof unless there is some significant difference from the case of S4LPN that is worth mentioning.

We prove proof confluence for TGLA, namely a version of the Church-Rosser property of a single-step prefixed tableau system (modulo renaming of prefixes). Massacci [102] briefly explains the notion as follows, "confluence means that the order in which we select the rules dose not substantially matter: we can always "converge" to the same result *without backtracking.*" (p.323 [102])

Now we give definitions ([82], [41]). Let us use x, y, z to stand for stages of the computation and $\longrightarrow$ to stand for a relation between them. In our context, the computation is the proof search in the prefixed tableau system, and its stages are tableaux. $x \longrightarrow y$ if the tableau y is obtained from x with an application of a tableau rule (a single step of reduction in the sense of a term rewriting systems). The relation $\longrightarrow^*$ is the reflexive and transitive closure of $\longrightarrow$, and $\longrightarrow^\epsilon$ is the reflexive closure. s and t are joinable (and we write $s \downarrow t$) if $\exists u(s \longrightarrow^* u$ and $t \longrightarrow^* u)$. s and t are meetable (and we write $s \uparrow t$) if $\exists r(r \longrightarrow^* s$ and $r \longrightarrow^* t)$.

Tableau rules are then formulated as the following reduction rules. (α, β, ν cases are the same as S4LPN.)

Reduction precondition	Reduction relation

π (GL) : $\sigma F\Box A \in \mathcal{B}$ and $\forall\Phi$, s.t. $\sigma.n\Phi \notin \mathcal{B}$ $\quad \mathcal{B} \longrightarrow_{\pi} \mathcal{B}\cup\{\sigma.nFA, \sigma.nT\Box A\}$.

π (Grz) : $\sigma F\boxdot A \in \mathcal{B}$ and $\forall\Phi$, s.t. $\sigma.n\Phi \notin \mathcal{B}$ $\quad \mathcal{B} \longrightarrow_{\pi} \mathcal{B}\cup\{\sigma.nFA,$
$\sigma.nT\boxdot(A \to \boxdot A)\}$.

We define an equivalence relation between stages of computation based on renaming of prefixes. In order to do that, let us first define renaming function h for prefixes ([102]) and then prove a proposition.

Definition 4.2.39 *An injective and surjective function h from the set of prefixes onto itself is a renaming if and only if $h(1) = 1$ and $h(\sigma.n) = h(\sigma).m$ for some integer m.*

Proposition 4.2.40 *A rule (r) can be applied to $\sigma\Phi$ in $\mathcal{B}_i$ if and only if it can be applied to $h_{ij}(\sigma)\Phi$ in $\mathcal{B}_j$.*

Proof For cases in which we do not change the prefix of the premise in applying the rule, applicability of a rule trivially holds. Since $\nu_{K,4}$ cases are the same as S4LPN (explicit versions with proof-terms are similar), we only check π-rules. Note that the domain of the mapping h is as follows: $Dom(h) = \{\sigma|\sigma\Phi \in \mathcal{B}_i$ for some Φ $\}$ where $\mathcal{B}_i$ is a branch in the original tableau. (Let the branch of its numerical variants be $\mathcal{B}_j$.)

π-rule (GL and Grz cases): $\sigma F\Box\varphi \in \mathcal{B}_i$ and $\mathcal{B}_i$ satisfies the precondition for applying π-rule, namely $\sigma.m\Phi \notin \mathcal{B}_i$ for any Φ. The first conjunct immediately implies $h(\sigma)F\Box\varphi \in \mathcal{B}_j$. To derive the desired second conjunct, suppose $h(\sigma.m)\Phi \in \mathcal{B}_j$ for some Φ. Then the mapping h has the image at

$\sigma.m$, i.e., $\sigma.m \in Dom(h)$. Thus, $\sigma.m\Phi \in \mathcal{B}_i$ for some Φ. But this contradicts the assumption that $\mathcal{B}_i$ satisfies the precondition for applying π-rule. The argument for the converse direction is similar. ($h_{ij}(\sigma)\Phi \in \mathcal{B}_j$ and the precondition is satisfied, i.e. $h_{ij}(\sigma).m\Phi \notin \mathcal{B}_j$ for any Φ. For some reductio, suppose $\sigma.m\Phi \in \mathcal{B}_i$ for some Φ. Then, $h_{ij}(\sigma.m)\Phi \in \mathcal{B}_j$ for some Φ. Hence, $h_{ij}(\sigma).m \in \mathcal{B}_j$. This contradicts the precondition.)

(Note that the neither $\sigma.mT\Box\varphi$ nor $\sigma.mT \boxdot (\varphi \to \boxdot\varphi)$ affects the argument.) ⊠

Then we can define an equivalence relation $\sim$ between stages of computation.

Definition 4.2.41 *The set of prefixed formulas* $\mathcal{B}_1$ *and* $\mathcal{B}_2$ *are equivalent modulo a renaming of prefixes if and only if there are two renamings* h_{12} *and* h_{21} *such that* $h_{ij}(h_{ji}(\sigma)) = \sigma$ *and if* $\sigma\Phi \in \mathcal{B}_i$ *then* $h_{ij}(\sigma)\Phi \in \mathcal{B}_j$ *for* $i, j = 1, 2$*.*

The definition can be extended to a tableau as sets of branches. Based on these definitions, we can state the definition of strong confluence and confluence.

Definition 4.2.42 *The relation* $\longrightarrow$ *is strongly confluent modulo* $\sim$ *if and only if* $\forall x_1, x_2, y_1, y_2$*, if* $x_1 \sim x_2$ *and* $x_1 \longrightarrow y_1$ *and* $x_2 \longrightarrow y_2$*, then* $\exists u_1, u_2$*, s.t.* $u_1 \sim u_2$ *and* $y_1 \longrightarrow^* u_1$ *and* $y_2 \longrightarrow^\epsilon u_2$*.*

Definition 4.2.43 (Confluence [41]) *A binary relation is confluent if any two elements are joinable when they are meetable* ($\uparrow \subseteq \downarrow$)

Theorem 4.2.44 (Strong Proof Confluence [82]) *If ν-formulas (also t : formulas in our case) can be reduced more than once, single step tableau rules are strongly confluent modulo isomorphic renaming of prefixes.*

Discussions of the content of the theorem and the details of the proof are almost the same as S4LPN. In the following, we omit the details except the cases in which details are somewhat different from S4LPN.

(It is worth mentioning that Reflection Rule is automatically excluded from our consideration about proof confluence since there is no choice in applying Reflection Rule. In addition, Reflection Rule has nothing to do with renaming either since $h(1) = 1$ always holds. However, this raises no problem in our translation, since there is no problem corresponding to Reflection Rule in HGLA. The translation between the two reflection rules in the two systems is straightforward.)

Massacci actually proves the following statement that is stronger than theorem 44 and derives the theorem as a corollary.

Lemma 4.2.45 (One-step confluence (modulo renaming)) $\forall x, x', y, z$, *if* $x \sim x'$ *and* $x \longrightarrow y$ *and* $x' \longrightarrow z$, *then* $\exists u, u_e$ ($y \longrightarrow u$ *and* $z \longrightarrow u_e$ *and* $u \sim u_e$).

Massacci uses what is called Knuth-Bendix method, which proves the critical pair lemma [102]. Massacci begins a proof of the lemma with an observation that formulas with different prefixes do not interact each other (whose proof is straightforward).

Proposition 4.2.46 *Let* $\sigma_1\Phi_1$ *and* $\sigma_2\Phi_2$ *be prefixed signed formulas with* $\sigma_1 \neq \sigma_2$ *and let* $\mathcal{B}'$ *be the reduction of* $\mathcal{B}$ *using rule* (r_1) *on* $\sigma_1\Phi_1$*. If rule* (r_2) *can be applied to* $\sigma_2\Phi_2$ *in* $\mathcal{B}$*, then it can be applied in* $\mathcal{B}'$*.*

Proof of the lemma We follow Massacci's presentation of a proof in giving the details of the case of the rule for K and a case where $x = x'$ and explaining how to extend it to more general cases.

Due to the proposition, the cases of superposition are reduced to the four cases: (1) $\sigma T\Box A$, $\sigma.n\Phi_1$ and $\sigma.m\Phi_2$ are on $\mathcal{B}$; (2) $\sigma T\Box A$, $\sigma T\Box C$ and $\sigma.n\Phi$ are on $\mathcal{B}$; (3) $\sigma T\Box A$, $\sigma F\Box C$ and $\sigma.n\Phi_1$ are on $\mathcal{B}$ but no prefixed formula $\sigma.m\Phi_2$ is present on $\mathcal{B}$ ($\sigma.m$ is new); (4) $\sigma F\Box A$, $\sigma F\Box C$ are on $\mathcal{B}$ and the prefixes $\sigma.n$ and $\sigma.m$ are new. (And we also have their explicit counterparts.)

(I) **Extension (or modification) of this proof to other logics**: These cases for K can be extended to other logics by replacing the prefixed formulas $\sigma T\Box A$, $\sigma.nTA$ for ν_K-rule with the premise and the conclusion of each ν-rule (and explicit rules). Strictly speaking, our π-rules here are different from that of K, so a case involving π-rule is not an extension but a modification. Also, each ν-formula (and t : formula) must be reducible more than once because each logic requires more than one ν-rule and all rules must be applicable. Due to formulas of the form $t : \varphi$, we have to check the following cases including (1)-(4).

We officially need to check all the four cases of superposition (1)-(4) for all the pertinent rules, and we did it. However, cases (1) and (2) are the same as S4LPN, and even other cases may be too tedious. So we show only some representative cases. (See section 4.1.4 about which rules (other than listed here) should be considered in which cases. Also, let us note that π-rule for GrzA can be handled similarly to that of GLA.)

Case (3) 3.1. $\sigma T\Box A$ (apply $\nu_{K,4}$ and, for GrzA, ν_T), $\sigma F\Box C$ and $\sigma.n\Phi_1$ are on $\mathcal{B}$ but no prefixed formula $\sigma.m\Phi_2$ is present on $\mathcal{B}$ ($\sigma.m$ is new).

3.2. $\sigma Tt : A$ (apply ET, EK, E4, EF, E4r) or $\sigma Ft : A$ (apply EF, EFr, or any operational rule), $\sigma F\Box C$ and $\sigma.n\Phi_1$ are on $\mathcal{B}$ (or $\sigma \neq 1$ for E4r or EFr, or no precondition for prefix in ET) but no formula $\sigma.m\Phi_2$ is present on $\mathcal{B}$ ($\sigma.m$ is new).

Proof of case (3): (Again we show only K.) We can use ν_K-rule or use π-rule and introduce a new prefix. If we use ν_K-rule, we do not introduce any new prefix and $\sigma.m$ would still be new in $\mathcal{B} \cup \{\sigma.nTA\}$. Thus, applying π-rule, we obtain $\mathcal{B} \cup \{\sigma.nTA, \sigma.mFC, \sigma.mT\Box C\}$. If we use π-rule first, then we obtain $\mathcal{B} \cup \{\sigma.mFC, \sigma.mT\Box C\}$. By assumption, for all Φ, $\sigma.m\Phi \notin \mathcal{B}$. Since $\sigma.nTD \in \mathcal{B}$, $\sigma.m \neq \sigma.n$. So, we can apply ν_K-rule and obtain $\mathcal{B} \cup \{\sigma.nTA, \sigma.mFC, \sigma.mT\Box C\}$.

Proof of case (4): (Cases are not increased here.) We must use renaming to prove confluence in this case. Suppose we reduce first $\sigma F\Box A$ and obtain $\mathcal{B} \cup \{\sigma.n_1FA, \sigma.n_1T\Box A\}$. In the new branch, the prefix $\sigma.n_1$ is no longer new. So, the next reduction forces the use of $\sigma.m_1$. Then we get $\mathcal{B}_1 = \mathcal{B} \cup \{\sigma.n_1 FA, \sigma.n_1 T\Box A,\ \sigma.m_1 FC, \sigma.m_1 T\Box C\}$. If we reduce first $\sigma F\Box C$ with $\sigma.n_2$ and obtain $\mathcal{B} \cup \{\sigma.n_2 FC,\ \sigma.n_2 T\Box C\}$. By another π-reduction, we get $\mathcal{B}_2 = \mathcal{B} \cup \{\sigma.n_2 FA, \sigma.n_2 T\Box A, \sigma.m_2 FC,\ \sigma.m_2 T\Box C\}$. At this stage, we can only guarantee that there are two new prefixes $\sigma.n$ and $\sigma.m$. It can be that $n_1 = n \neq m = n_2$ and $A \neq B$. Thus, $\mathcal{B}_1 \neq \mathcal{B}_2$. Then we define two renamings h_{12} and h_{21}.

$$h_{ij}(s) = \begin{cases} \sigma.n_j & \text{if } s = \sigma.n_i \\ \sigma.m_j & \text{if } s = \sigma.m_i \\ \text{s} & \text{otherwise} \end{cases}$$

Now we can prove 1. $\sigma\Phi \in \mathcal{B}_i \Longrightarrow h_{ij}(\sigma)\Phi \in \mathcal{B}_j$ and 2. $h_{ij}(h_{ji}(\sigma)) = \sigma$.

(II) **Extension of the proof to a general case where $x \sim x'$**

The proof of the general case $x \sim x'$ uses more complicated notations but follows the same pattern, since we have two branches $\mathcal{B}_1 \sim \mathcal{B}_2$ and two renamings $h_{12} : \mathcal{B}_1 \longrightarrow \mathcal{B}_2$ and $h_{21} : \mathcal{B}_2 \longrightarrow \mathcal{B}_1$. The proposition for the general case is as follows.

Proposition 4.2.47 *Let $\sigma_1\Phi_1$ and $\sigma_2\Phi_2$ be prefixed signed formulas such that $h_{12}(\sigma_1) \neq \sigma_2$ and let $\mathcal{B}_1^r$ be the reduction of $\mathcal{B}_1$ using rule (r) on $\sigma_1\Phi_1$ and let $\mathcal{B}_2^r$ be the reduction of $\mathcal{B}_2$ using rule (r) on $h_{12}(\sigma_1)\Phi_1$. If rule (r_2) can be applied to $\sigma_2\Phi_2$ in $\mathcal{B}_2$, then it can be applied in $\mathcal{B}_2^r$.*

The argument for the proposition is essentially the same as that for S4LPN except for the presence of $\sigma.nT\Box\varphi$ (π-rule for GL) and $\sigma.nT \boxdot (A \to \boxdot A)$ (π-rule for Grz). As in the previous case (3), (4), these formulas do not affect the structure of the argument.

4.2.3.2.4. Numerical variants of closed tableaux

Like the case of S4LPN, to eliminate the case ii) 2) in the subsection b, we prove a general proposition stating the closure of a "numerical variant" of a closed tableau in a prefixed tableau system. By "a numerical variant of a prefixed formula," we mean a prefixed formula that is different only in the number of the prefix from the formula. This can be taken as constructing a case of Massacci's renaming function defined in the previous section. We call a prefix $h(\sigma)$ a numerical variant of σ if there is at least one i-th element of σ s.t. $h(n_i) \neq n_i$. As in the case of S4LPN, we can prove the following.

Proposition 4.2.48 *If a tableau $\mathcal{T}_i$ is closed, then its numerical variant $\mathcal{T}_j$ is also closed, provided that it satisfies the preconditions of applications for π, ν, and E rules.*

Proof Similar to S4LP. $\boxtimes$

Now we apply the general idea of handling numerical variants of closed tableaux to the aforementioned combination of E4r (or EFr) and $\nu_{K,4}$. Here also, we call a "reuse of prefix" an application of $\nu_{K,4}$ on a prefix that has already been used for applying reverse rules. We prove that such applications of $\nu_{K,4}$ can be eliminated from a proof with such applications.

Lemma 4.2.49 *If a tableau is closed by using reuses of prefixes by $\nu_{K,4}$-rules, then a closed tableau can be constructed without using such applications.*

Proof The proof is similar to the case of $\mathsf{S4LPN}$ (again except the treatment of extra formulas $\sigma.nT\Box\varphi$ (GL) and $\sigma.nT\boxdot(\varphi \to \boxdot\varphi)$). But obviously these do not disturb the argument. ⊠

E.g. The example of a tableau proof given above is modified as follows.

$1F\Box(t : \Box A \to A)$

$1.1.Ft : \Box A \to A$; $1.1T\Box(t : \Box A \to A)$

$1.1.Tt : \Box A$; $1.1.FA$

$1Tt : \Box A$; $1T\Box A$

$\mathbf{1.2.Ft : \Box A \to A}$; $\mathbf{1.2T\Box(t : \Box A \to A)}$

$\mathbf{1.2.Tt : \Box A}$; $\mathbf{1.2.FA}$

$\mathbf{1Tt : \Box A}$; $\mathbf{1T\Box A}$

$1.2T\Box A$

$1.2TA$

×

4.2.3.2.5. Prefixed tableau proofs translatable to hypersequent proofs

Having these preparations, we are ready to transform any given prefixed tableau proof into a prefixed tableau proof translatable to a hypersequent proof. A normal form of a proof in the prefixed tableau system is defined as before, and we prove a proposition.

Definition 4.2.50 *We call a prefixed tableau proof in a π-$\nu_{K,4}$ normal form when for any application of $\nu_{K,4}$, there exists an application of π such that the prefix used in the application of $\nu_{K,4}$ is introduced by the application of π and there exists no application of rule x such that $\pi \gg x \gg \nu_{K,4}$ (except other applications of $\nu_{K,4}$).*

Proposition 4.2.51 *Suppose there exists a tableau proof of φ in one of the prefixed tableau systems TGLA and TGrzA. Then, this tableau proof can be effectively transformed into a tableau proof that is in a π-$\nu_{K,4}$ normal form.*

Proof The proof is similar to the case of S4LPN. ⊠

E.g. The example given above is not yet in a π-$\nu_{K,4}$ normal form. By appealing to proof confluence, the tableau rules can be rearranged and modified into a π-$\nu_{K,4}$ normal form as follows.

$1F\Box(t : \Box A \to A)$

$1.1Ft : \Box A \to A$; $1.1T\Box(t : \Box A \to A)$

$1.1Tt : \Box A$; $1.1FA$

$1Tt : \Box A$

$1T\Box A$

$\mathbf{1.2Ft : \Box A \rightarrow A}$; $\mathbf{1.2T\Box(t : \Box A \rightarrow A)}$

$1.2T\Box A$; $1.2TA$

$\mathbf{1.2Tt : \Box A}$; $\mathbf{1.2FA}$

×

Once we obtain a π-$\nu_{K,4}$ normal form, we replace all the applications of π by $\pi^\sharp$. The original proof is still a proof but then application of $\nu_{K,4}$ rules become redundant by this modification of the proof.

4.2.3.2.6. Inductive proof of preserving provability under the translation

We introduce auxiliary prefixed tableau systems $\mathsf{TL}^\sharp$, where L stands for GLA, or GrzA. In $\mathsf{TL}^\sharp$, instead of using π-rule, we use $\pi^\sharp$-rule so that we can handle π-rule and $\nu_{K,4}$ simultaneously. (Here we mainly discuss GLA, but the pattern of the argument is common in both of these cases. We show only some crucial cases of GrzA later.)

Prefixed tableau systems $\mathsf{TL}^\sharp$ is useful for handling the discrepancies between prefixed tableau systems and hypersequent calculi, but they bring in some complication to the induction on the depth of a tableau proof due to the existence of $\nu_{K,4}$ in $\mathsf{TL}^\sharp$. To make the inductive proof simpler, we use a subsystem $\mathsf{TL}^{\sharp\circ}$ of $\mathsf{TL}^{\sharp}$ in which we keep $\pi^{\sharp}$ rule, but we remove $\nu_{K,4}$ without weakening the deductive power of $\mathsf{TL}^{\sharp}$. Namely, we treat $\nu_{K,4}$ as admissible rules in $\mathsf{TL}^{\sharp\circ}$, since due to the theorem stating the existence of π-$\nu_{K,4}$ normal form, we can show that $\nu_{K,4}$ can be absorbed into $\pi^{\sharp}$-rule. Our translation to the hypersequent calculi is eventually defined from $\mathsf{TL}^{\sharp\circ}$ to the hypersequent calculi. We claim that due to the theorem stating the existence of π-$\nu_{K,4}$ normal form, we can show that $\nu_{K,4}$ can be absorbed into $\pi^{\sharp}$-rule.

Proposition 4.2.52 *1. In TL, where $\mathsf{L}\in\{\mathsf{GLA}, \mathsf{GrzA}\}$, $\pi^{\sharp}$ is a derived rule. 2. For any φ, if φ has a tableau proof in TL, then φ has a tableau proof in $\mathsf{TL}^{\sharp\circ}$.*

Proof Again, except the extra formulas for π-rules ($\sigma.nT\Box\varphi$ and $\sigma.nT\boxdot(\varphi \to \boxdot\varphi)$, the proof is similar to that of $\mathsf{S4LPN}$. ⊠

Lemma 4.2.53 *Let S be a finite set of prefixed signed formulas. If there is a closed tableau for S using the prefixed tableau for GLA, then the hypersequent S^s is provable in HGLA^{-}.*

Proof Before going into the inductive proof, we go through the following process to make sure that the proof we translate into a proof in the hypersequent calculus.

Suppose that there is a closed tableau in TGLA for S. Fix one. We call it $\mathcal{T}_1$. By using the manipulations of prefixed tableau proofs given in the previous section, we transform $\mathcal{T}_1$ into a prefixed tableau proof in a π-$\nu_{K,4}$ normal form. (Note that this normal form is not unique.) Call the proof $\mathcal{T}_1^{*}$.

Modify $\mathcal{T}_1^*$ into $(\mathcal{T}_1^*)^\sharp$ by replacing all applications of π by $\pi^\sharp$ (This may produce some extra applications of $\nu_{K,4}$ that the original proof did not have. Hence, our translation has these redundancies. However, it is obvious that this does not affect the provability.) Since all the applications of $\nu_{K,4}$ are absorbed in $\pi^\sharp$, we can take the resulting proof as a proof in $\mathsf{TGLA}^{\sharp\circ}$, i.e. $(\mathcal{T}_1^*)^{\sharp\circ}$.

Then apply the mapping s from the prefixed tableau $(\mathcal{T}_1^*)^{\sharp\circ}$ (= a set of prefixed formulas) to a hypersequent $((\mathcal{T}_1^*)^{\sharp\circ})^s$. (The mapping is not about inference rules, but this is from the deductive meta-language of TGLA to the deductive meta-language of HGLA).

Now we show by induction on the depth d of a prefixed tableau in the system $\mathsf{TGLA}^{\sharp\circ}$ that for any $(\mathcal{T}_1^*)^{\sharp\circ}$, $((\mathcal{T}_1^*)^{\sharp\circ})^s$ is provable in HGLA^-. Here the depth is a number d such that there is a closed prefixed tableau in $\mathsf{TGLA}^{\sharp\circ}$ for S with d applications of tableau rules. Suppose a tableau for S closes with depth d ($d \geq 0$) and, by IH, the theorem holds for sets that close with less than d. Suppose we have made the first application of a tableau rule. Then, to obtain a closed tableau, we only need $d-1$ applications of the rules. For any tableau rule, after this first application we apply IH, since now we have only depth $d-1$. (Note that $d-1$ makes sense only for $d \leq 1$. IH can be applied only for $d \leq 1$. We take $d = 0$ is a base case in which the given set is already a closed tableau without applying any rule. Also, note that any closed tableau can be translated into a derivable hypersequent in HGLA^- by its axioms, IW and EW.) By IH, there must be a proof of a hypersequent obtained by using the mapping s to the tableau resulting by applying one of the tableau rules. Hence, the remaining step is to show that we can prove the hypersequent obtained by taking an image of the mapping s of the set of the prefixed signed formulas we had before we apply the rule. We treat only those rules that are different from rules of S4LPN.

Case 1. Reflection Rule:

Suppose $\{1F\varphi\} = S$. Note that due to the restricted nature of this rule, we have only a singleton (with the prefix 1) as the conclusion. By the assumption, such an S has a closed tableau with depth d. Since the first application of a rule in this tableau is Reflection Rule, $\{1F\Box\varphi\}$ has a closed tableau with depth $d-1$. Due to the destructive feature of Reflection Rule, the conclusion of Reflection Rule in TGLA is also a singleton. Thus, when we turn "upside down" (reverse the order of an application of) the tableau rule, we have the singleton $\{1F\Box\varphi\}$ as the premise.

By IH, $(\{1F\Box\varphi\})^s$ is provable in HGLA$^-$. By definition of s, $\{1F\Box\varphi\}^s$ is given as $\Rightarrow \Box\varphi$. By using Reflection Rule in HGLA$^-$, it is clear that we can derive $\Rightarrow \varphi$.

Case 2. $\pi^\sharp$-rules for GL

$\sigma F\Box\varphi \in S$, and $S^\sharp_{\sigma.n} = \{\sigma.nT\Box\psi, \sigma.nT\psi | \sigma T\Box\psi \in S\}$. So, $(S^\sharp_{\sigma.n})^s = \Gamma^\sharp_{\sigma.n}$. Note that $\{\Box\psi|\Box\psi \in \Gamma^\sharp_{\sigma.n}\} \subseteq \Gamma_\sigma$. Then by IH, $(S \cup \{\sigma.nF\varphi\} \cup \{\sigma.nT\Box\varphi\} \cup S^\sharp_{\sigma.n})^s$, which is identical with $\Gamma_1 \Rightarrow \Delta_1| \dots |\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\Box\varphi, \Gamma^\sharp_{\sigma.n} \Rightarrow \varphi| \dots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$, is provable in HGLA$^-$. We can derive S^s from the above hypersequent in HGLA$^-$ as follows.

$$\dfrac{\dfrac{\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\Box\varphi, \Gamma^\sharp_{\sigma.n} \Rightarrow \varphi|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\Box\varphi, \Box\psi_1, \psi_1, \dots, \Box\psi_k, \psi_k \Rightarrow \varphi|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\Box\psi_1, \dots, \Box\psi_k \Rightarrow \Box\varphi|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\;\Box\mathsf{GL}}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\;\mathrm{IW}}{\Gamma_1 \Rightarrow \Delta_1|\dots|\Gamma_\sigma \Rightarrow \Delta_\sigma, \Box\varphi|\dots|\Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\;(\mathrm{set})$$

This completes all the cases that we need to check here. ⊠ (Lemma)

Note: 1. The lemma is stated with a set of prefixed formula S. A closed tableau in TGLA constitutes a "tableau proof" of a formula φ if and only if it start with a singleton of a prefixed signed formula $1F\Box\varphi$.

2. Due to the presence of one extra $\Box$ in the definition of a tableau proof of φ in the case of TGLA, we always need one extra step of Reflection Rule in HGLA in order to prove *the same* formula as the one that has a tableau proof.

3. Due to the proposition 4.5.16, we can work in $\mathsf{TGLA}^{\sharp\circ}$. Therefore, we do not have $\nu_{K,4}$ rules as primitive rules in $\mathsf{TGLA}^{\sharp\circ}$.

E.g. The example given above in a $\pi - \nu_{K,4}$ normal form can be translated into the following hypersequent proof. (Let us use an abbreviation $\Phi := t : \Box A \to A$ in the following derivation.)

$$
\begin{array}{cl}
A \Rightarrow A & \\
\cline{1-1}
{*}\ t:\Box A, \Box A \Rightarrow \Box\Phi | \Box\Phi, t:\Box A \Rightarrow \Phi, A | \Box\Phi, t:\Box A, \Box A, A \Rightarrow \Phi, A & \text{IW, EW} \\
\cline{1-1}
{*}\ t:\Box A, \Box A \Rightarrow \Box\Phi | \Box\Phi, t:\Box A \Rightarrow \Phi, A | \Box\Phi, \Box A, A \Rightarrow t:\Box A \to A & \\
\cline{1-1}
{*}\ t:\Box A, \Box A \Rightarrow \Box\Phi | \Box\Phi, t:\Box A \Rightarrow \Phi, A | \Box\Phi, \Box A, A \Rightarrow \Phi & \\
\cline{1-1}
{*}\ t:\Box A, \Box A \Rightarrow \Box\Phi | \Box\Phi, t:\Box A \Rightarrow \Phi, A | \Box A \Rightarrow \Box\Phi & \Box\text{GL} \\
\cline{1-1}
t:\Box A, \Box A \Rightarrow \Box\Phi | \Box\Phi, t:\Box A \Rightarrow \Phi, A | t:\Box A, \Box A \Rightarrow \Box\Phi & \text{IW} \\
\cline{1-1}
{*}\ t:\Box A, \Box A \Rightarrow \Box\Phi | \Box\Phi, t:\Box A \Rightarrow \Phi, A & \text{(set)} \\
\cline{1-1}
t:\Box A, t:\Box A \Rightarrow \Box\Phi | \Box\Phi, t:\Box A \Rightarrow \Phi, A & \text{Lt:} \\
\cline{1-1}
{*}\ t:\Box A \Rightarrow \Box\Phi | \Box\Phi, t:\Box A \Rightarrow \Phi, A & \text{(set)} \\
\cline{1-1}
t:\Box A \Rightarrow | \Rightarrow \Box\Phi | \Box\Phi, t:\Box A \Rightarrow \Phi, A & \\
\cline{1-1}
\Box\Phi, t:\Box A \Rightarrow \Phi, A | \Rightarrow \Box\Phi | \Box\Phi, t:\Box A \Rightarrow t:\Box A \to A, A & \text{IW} \\
\cline{1-1}
{*} \Rightarrow \Box\Phi | \Box\Phi, t:\Box A \Rightarrow \Phi, A & \text{(set)} \\
\cline{1-1}
\Rightarrow \Box\Phi | \Box\Phi \Rightarrow \Phi & \text{R}\to\text{, (set)} \\
\cline{1-1}
\Rightarrow \Box\Phi | \Rightarrow \Box\Phi & \\
\cline{1-1}
{*} \Rightarrow \Box(t:\Box A \to A) & \\
\cline{1-1}
\Rightarrow t:\Box A \to A & \text{Reflection}
\end{array}
$$

Note that lines with * have an exactly corresponding step in the given tableau proof. Intermediate steps are inserted to complete hypersequent proofs (some trivial steps are omitted from presentation, since they make the proof excessively tedious). The last is needed since the prefixed tableau system uses an extra $\Box$.

Although in many cases we have to insert some intermediate lines in hypersequent proofs, hypersequent proofs are powerful enough to reconstruct exactly corresponding lines in the given prefixed tableau proofs in a stepwise manner.

We can prove the case of HGrzA in a similar way. Most cases are the same as HGLA, so we will show only ν_T-rule and π-rule for $\boxdot$.

Lemma 4.2.54 *Let S be a finite set of prefixed signed formulas. If there is a closed tableau for S using the prefixed tableau for GrzA, then the hypersequent S^s is provable in HGrzA.*

Proof Here we also use the auxiliary system TGrzA$^{\sharp\circ}$ to show preservation of derivability under the mapping s. The details are similar to both of the cases of S4LPN and GLA. (Since we do not have the special condition for HGLA, it is even closer to S4LPN. Namely, since we do not need any extra $\Box$ for the definition of a formula having a tableau proof in TGrzA, we do not need the remark concerning the extra $\Box$ made in the case of TGLA.)

Now we prove the theorem by induction of the depth of tableau proofs.

Case 1. ν_T-rule (for GrzA).

Suppose $\sigma T \boxdot \varphi \in S$. By the assumption, such an S has a closed tableau with depth d. Since the first application of a rule in this tableau is T-rule, $S \cup \{\sigma T \varphi\}$ has a closed tableau with depth $d-1$. By IH, $(S \cup \{\sigma T \varphi\})^s$ is provable in TGrzA$^{\sharp\circ}$. By definition of s, S^s is as follows: $\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma, \boxdot\varphi \Rightarrow \Delta_\sigma | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$. Then we have the following derivation, which shows provability of S^s in HGrzA.

$$\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma, \boxdot\varphi, \varphi \Rightarrow \Delta_\sigma | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}{\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma, \boxdot\varphi, \boxdot\varphi \Rightarrow \Delta_\sigma | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\,\mathrm{L}\boxdot}{\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma, \boxdot\varphi \Rightarrow \Delta_\sigma | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}}\,(\mathrm{set})$$

Case 2. $\pi^\sharp$-rule for Grz

$\sigma F \boxdot \varphi \in S$, and $S^\sharp_{\sigma.n} = \{\sigma.nT\boxdot\psi, \sigma.nT\psi | \sigma T \boxdot \psi \in S\}$. So, $(S^\sharp_{\sigma.n}) = \Gamma^\sharp_{\sigma.n}$. Note that $\{\boxdot\psi | \boxdot\psi \in \Gamma^\sharp_{\sigma.n}\} \subseteq \Gamma_\sigma$. Then by IH, $(S \cup \{\sigma.nF\varphi\} \cup \{\sigma.nT\boxdot(\varphi \to \boxdot\varphi)\} \cup S^\sharp_{\sigma.n})^s$, which is identical to $\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma \Rightarrow \Delta_\sigma, \boxdot\varphi | \boxdot(\varphi \to \boxdot\varphi), \Gamma^\sharp_{\sigma.n} \Rightarrow \varphi | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m}$, is provable in HGrzA^-. We can derive S^s from the above hypersequent in HGrzA^- as follows. (Let us use the notation $\overrightarrow{\boxdot\psi_i, \psi'_i}$ for $\boxdot\psi_1, \psi_1 \ldots, \boxdot\psi_k, \psi_k$ and $\overrightarrow{\boxdot\psi'_i}$ for $\boxdot\psi_1, \ldots, \boxdot\psi_k$.)

$$\begin{array}{cl}
\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma \Rightarrow \Delta_\sigma, \boxdot\varphi | \boxdot(\varphi \to \boxdot\varphi), \Gamma^\sharp_{\sigma.n} \Rightarrow \varphi | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} & \\ \hline
\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma \Rightarrow \Delta_\sigma, \boxdot\varphi | \boxdot(\varphi \to \boxdot\varphi), \overrightarrow{\boxdot\psi_i, \psi'_i} \Rightarrow \varphi | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} & \nu_T \\ \hline
\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma \Rightarrow \Delta_\sigma, \boxdot\varphi | \boxdot(\varphi \to \boxdot\varphi), \overrightarrow{\boxdot\psi_i, \boxdot\psi'_i} \Rightarrow \varphi | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} & \text{(set)} \\ \hline
\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma \Rightarrow \Delta_\sigma, \boxdot\varphi | \boxdot(\varphi \to \boxdot\varphi), \overrightarrow{\boxdot\psi'_i} \Rightarrow \varphi | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} & \boxdot\mathsf{Grz} \\ \hline
\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma \Rightarrow \Delta_\sigma, \boxdot\varphi | \overrightarrow{\boxdot\psi'_i} \Rightarrow \boxdot\varphi | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} & \text{IW} \\ \hline
\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma \Rightarrow \Delta_\sigma, \boxdot\varphi | \Gamma_\sigma \Rightarrow \Delta_\sigma, \boxdot\varphi | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} & \text{(set)} \\ \hline
\Gamma_1 \Rightarrow \Delta_1 | \ldots | \Gamma_\sigma \Rightarrow \Delta_\sigma, \boxdot\varphi | \ldots | \Gamma_{\sigma_m} \Rightarrow \Delta_{\sigma_m} &
\end{array}$$

This completes all the cases that we need to check here. ⊠

Note: We do not have $\nu_{K,4}$ rules as primitive rules in $\mathsf{TGrzA}^{\sharp\circ}$, either.

The completeness theorems for HGLA and HGrzA immediately follows from the above theorems as corollaries. Here, again, HGLA^- (HGrzA^-) stands for HGLA (HGrzA) without cut.

Theorem 4.2.55 (Weak Completeness)

1. If φ is valid in the semantics for GLA, then $\mathsf{HGLA}^- \vdash \Rightarrow \varphi$.

2. If φ is valid in the semantics for GrzA, then $\mathsf{HGrzA}^- \vdash \Rightarrow \varphi$.

Proof 1. Suppose φ is valid in the semantics of GLA. Then we can construct a tableau proof of φ in TGLA. Then we can translate the tableau proof of φ to a cut-free proof in the hypersequent calculus HGLA.

2. This case is similar to 1. ⊠

Cut-admissibility for HGLA and HGrzA follows almost immediately from the theorem via a well-known method combining soundness of the system with cut and completeness without cut.

Corollary 4.2.56 (Cut-admissibility) *1. If $\mathsf{HGLA} \vdash \Rightarrow \varphi$, then $\mathsf{HGLA}^- \vdash \Rightarrow \varphi$.*

2. If $\mathsf{HGrzA} \vdash \Rightarrow \varphi$, then $\mathsf{HGrzA}^- \vdash \Rightarrow \varphi$

Proof Similar to the case of HS4LPN. ⊠

As immediate corollaries of cut-admissibility, we can show some statements.

Let us give some definitions first.

A theory T_2 is called a conservative extension of T_1 if the following holds.

1) T_2 is formulated in the language L_2 that includes the language L_1 in which T_1 is formulated.

2) For any formula $\varphi \in L_1$, if $T_2 \vdash \varphi$, then $T_1 \vdash \varphi$.

More informally, a theory T_2 is a conservative extension of another theory T_2 if any formula in a smaller language that formulates T_1 can be proven in T_2, it can be already proven in T_1.

Also, "the modal disjunction property" (MDP) refers to the following property in modal logics:

(MDP) If $L \vdash \Box\varphi \vee \Box\psi$, then $L \vdash \varphi$ or $L \vdash \psi$.

Proposition 4.2.57 *1. GrzA is a conservative extension of Grz*

2. GLA is a conservative extension of GL

Proof Observe that once a proof-term occurs in a cut-free proof in HGLA or in HGrzA, it never disappear. ⊠

Note that not only logic of proofs part does not affect purely modal part (GL) of GLA, but reflection rule is admissible in GL; hence, it does not increase the set of theorems in GL.

Proposition 4.2.58 *For any formula φ, ψ that does not contain any proof-term, if $L\vdash \Box\varphi \vee \Box\psi$, then $\vdash \varphi$ or $\vdash \psi$, where $L = \{$GrzA, GLA$\}$.*

Proof This is an immediate consequences of the above proposition and MDP for GL and Grz. ⊠

Note that it is well-known that modal disjunction property holds for GL and Grz themselves. The proposition immediately follows from the conservativeness result and the modal disjunction property of GL and Grz.

There is another corollary of cut-admissibility, which is an extension of the disjunction property of logic of proofs.

Proposition 4.2.59 *If* $L \vdash \Rightarrow t : \varphi \vee s : \psi$, *then* $L \vdash t : \varphi$ *or* $L \vdash s : \psi$, *for* $L \in \{\mathsf{GrzA}, \mathsf{GLA}\}$.

Proof The proof is similar to the case of S4LPN. ⊠

4.3 Discussion

Before closing, we would like to discuss a few further issues related to GLA and GrzA.

1. As we discussed in the introduction, there is a variant of GLA, namely LPP ([178]). The difference is that LPP has additional proof operators producing a new proof term connecting a proof term and a modal operator $\Box$. Namely, we do have the following two operators in the language: $\Uparrow^t_\Box$, $\Downarrow^\Box_t$. The axioms governing these new operators in LPP are as follows.

1. $t : \Box\varphi \rightarrow \Downarrow^\Box_t : \varphi$ 2. $t : \varphi \rightarrow \Uparrow^t_\Box : \Box\varphi$

One of the reasons why such term forming operators are desirable in the system is that using these devices, we can prove the following theorem.

Theorem 4.3.1 (Constructive necessitation) *If* $\mathsf{L}\vdash \varphi$*, then* $\mathsf{L}\vdash t:\varphi$ *for some proof-term* t*, where* $\mathsf{L}\in \{\mathsf{LPP}, \mathsf{LPGL}\}$.

Yavorskaya [178] proves this theorem by explicitly using these term-forming operators. On the other hand, Artemov and Nogina [7] does not have any such extra items in the language. However, they also prove constructive necessitation in their work. It turns out that if we use the axiom $t:\Box\varphi \to \varphi$, then the operators that Yavorskaya introduced are not necessary as primitive items in the language to prove constructive necessitation. In order to show this in LPGL, the axiom $t:\Box\varphi\to\varphi$ postulated in [7] play a crucial role. They point out that the axiom is, in a way, redundant,[48] since the axiom is derivable even in the subsystem of LPGL with empty constant specification. (Also, note that this axiom can easily be derived in LPP.) However, the axiom $t:\Box\varphi \to \varphi$ makes it possible to prove constructive necessitation without using $\Downarrow_t^{\Box}$, since with the axiom $t:\Box\varphi\to\varphi$ a term playing the role of $\Downarrow_t^{\Box}$ can be explicitly constructed in the language of LPGL. Hence, in a way, the issue of whether we introduce $\Downarrow_t^{\Box}$ or introduce the axiom $t:\Box\varphi\to\varphi$ is a matter of whether we extend the language or the axioms.

Even though these new items do not seem really necessary to have a satisfactory system for GLA anymore, it may be still an interesting question

---

[48] Reflection rule is also in a way redundant, since it is not only admissible like in GL but derivable from the rest of LPGL. However, reflection rule is needed to to guarantee a good behavior of fragments of LPGL with specific constants specifications ([7]).

whether or not we can handle these items by using our proof systems. It turns out to be straightforward to add these items in our prefixed tableau system and hypersequent calculus HGLA. We need the following additional rules for prefixed tableau system and the hypersequent calculus.

$$\frac{\sigma F \Downarrow_t^{\Box} : \varphi}{\sigma F t : \Box\varphi} \qquad \frac{\sigma F \Uparrow_{\Box}^{t} : \Box\varphi}{\sigma F t : \varphi}$$

$$\frac{G|\Gamma \Rightarrow \Delta, t : \Box\varphi}{G|\Gamma \Rightarrow \Delta, \Downarrow_t^{\Box} : \varphi} \qquad \frac{G|\Gamma \Rightarrow \Delta, t : A}{G|\Gamma \Rightarrow \Delta, \Uparrow_{\Box}^{t} : \Box\varphi}$$

To handle these proof operations in our setting, it suffices to modify closure conditions of evidence function in Fitting-style semantics. In particular, we add the following closure conditions to the one given above. (Here $u \in K$ and t is a proof term.)

7. $\Box\varphi \in \mathcal{E}(u, t) \Rightarrow \varphi \in \mathcal{E}(u, \Downarrow_t^{\Box})$
8. $\varphi \in \mathcal{E}(u, t) \Rightarrow \Box\varphi \in \mathcal{E}(u, \Uparrow_{\Box}^{t})$

Proof of soundness and completeness for the prefixed tableau system can be modified accordingly in a straightforward manner. Also, the translation from the prefixed tableaux to hypersequents can be extended to handle these cases straightforwardly.

2. Although our systems TGLA and HGLA are cut-free, there are items

in these systems that hinder satisfying the subformula property. As briefly mentioned in a footnote of the introductory subsection, one is about the application rule in LP and the other is the Reflection Rule. However, compared with the first case, the violation of the subformula property by the latter is not very harmful. There is a positive reason to think that this case is benign. It is obvious from our construction of a model in the proof of completeness that the only formulas that are outside of the subformula of the formula to be proven (call it φ) are the ones that have extra $\Box$'s in front of φ. Since our system is cut-free, this case is more like the analytic superformula property that modal logicians sometimes talk about [75], since we have a definite bound of the formulas to think about in proof search (namely, it suffices if we add at most n-many $\Box$ where n is the number of $\Box$-ed subformula in the formula to be proven). The case of application rule in LP is a deeper problem because there is no bound of the set of formulas to check in proof search.

3. The last thing to be discussed here is related to the first case of the violations of the subformula property mentioned above. There are at least several things yet to be done in proof theory of the logic of proofs and provability. Among these, we point out just two. A relatively minor one is concerned with a possibility of proving cut-elimination by using a purely syntactic reduction method. We could have started considering purely syntactic proof of cut-elimination in these cases, but there are a reason why that we did not. The reason is that a purely syntactic reduction method is already tricky for GL itself ([77]). Also, we thought that we probably should

first pursue the subformula property.

The other problem is highly relevant to this issue of why the subformula property is difficult to obtain in our system. One of the major open problems in the logic of proofs (or justification logics) is to discover natural Gentzen-style sequent calculi that enjoy the subformula property (in particular, handling $\cdot$-rule in some way). Related to this issue, recently G. Jäger introduced an interesting terminology, an internal cut and an external cut, in his survey article [88]. An external cut is just an ordinary cut. Jäger pointed out that although the logic of proofs enjoys cut-elimination, the current version of cut-elimination is of a limited nature, since we have proof application operation in proof-terms and the axiom that governs this feature of the logic $t : (\varphi \to \psi) \to (s : \varphi \to t \cdot s : \psi)$ can be taken as an internal cut. What is problematic in this cut-elimination is that it does not imply the subformula property as is obvious from the form the application axiom and the corresponding operational rule in sequent calculus for LP. The current approach in the logic of proofs does not have a perfectly unproblematic method of eliminating this operation. Although in many senses having a proof-term with $\cdot$ is better than an ordinary case of cut (in particular, we have some trace of which formulas were involved in a derivation of the pertinent formula by means of internalized modus ponens), it is true that the lack of the subformula property rests on this. In this sense, elimination of an internal cut is an interesting theoretical goal in this field. However, this seems to require a fundamental reconsideration of the basic format of LP, which currently uses the style of typed combinatory logic. Discovering a satisfactory proof system for logic of proofs with the subformula property may need a drastic conceptual innovation and paves a further deepening of the understanding of the nature of proofs along the line of logic of proofs. But it seems to be a very challenging problem.

Chapter 5

Conclusion

In this thesis, we have (philosophically) argued for one main thesis of the philosophical part of the thesis. In the technical part of the thesis, we proved cut-admissibility for several proof systems in modal and justification logics.

The main thesis of the philosophical part (Chapter 2) states that some modal operators are decent logical constants. We have done this by formulating our own structural-reflective view of logic, according to which logical constants are taken as results of reflecting operations in the metalanguage that is fixed for various different logics and the differences of these logics should be understood as structural differences. In Chapter 2, we also discussed the issue of proof-theoretic semantics. We argued that our view is not only coherent with at least one notion of meaning discussed in the literature of proof-theoretic semantics, i.e., operational meaning, but also the formulation of definition can be taken to be a generalization of an important concept (intrinsic) harmony in proof-theoretic semantics. We also discussed that fundamental notions in logic such as conservativeness and uniqueness and have concluded that the philosophical implications that are traditionally given to these notions at least as criteria for logical constant-hood have been overestimated. Although there is no doubt that these notions still play some role in philosophy of logic, we should also start thinking about some other approaches in pursuing the foundations of logic(s). For instance, the notion of "horizontalization" introduced by Dana Scott has not attracted very much attention, but some metatheoretic notions like this may deserve more investigations, although in this thesis we did not go too far into that direction. Also, we argued for the importance of hypersequent calculi there as follows. Although hypersequent may have limitations in formulating modal logics in a proof-theoretically satisfactory manner (especially in a cut-free manner) in the specific sense that the modal logics that can be covered by hypersequent calculi are more limited compared with other generalizations of sequent calculi, hypersequent calculi can still handle the kind of modal logics that are typically considered to be "logical" according to our philosophical view presented in the chapter. In addition, we gave a defense of the framework of hypersequent calculi by arguing that they have the virtue of simplicity and are less radically different from Gentzen's original sequent calculi from a methodological point of view.

On the other hand, the technical results that we have proved can be summarized as follows. In Chapter 3, we proved the cut-elimination theorems for hypersequent calculi for strict implication logics (SIS4, SIS4.3, SIS5) and modal logics (S4, S4.2, S4.3, S5 and S4+L in which L stands for S4.2, S4.3, S5) by using a syntactic method, called the substitution method. In the first half of Chapter 4, we semantically proved cut-admissibility of prefixed tableau systems for S4LPN and S4LP, and we proved cut-admissibility of hypersequent calculi for S4LP and S4LPN. In the latter half of Chapter 4, we proved cut-admissibility for prefixed tableau systems for GrzA and GLA, and we also proved cut-admissibility of hypersequent calculi for GrzA and GLA.

Since we have already extensively discussed the significance of these results that use hypersequent calculi in Chapter 2, we will not go into such a discussion here. However, it would be fair to say that we have extended the area of applications of hypersequent calculi to modal logics and justification logics in a non-trivial manner, although these results, in particular the ones given in Chapter 3, may not be too surprising due to the existence of cut-free hypersequent calculi for superintuitionistic counterpart of S4.2 and S4.3.

However, it is probably worth mentioning that the cutting-line between the modal logics for which the uniform syntactic method of substitution for cut-elimination works and the logics for which the method does not work almost coincides with the cutting-line between the logics that deserve being called "logics" from our structural-reflective point of view (e.g., S4.3) and the logics that are traditionally called logics but conceptually that go beyond our notion of "logic" for the point of view (e.g., S4.3Grz or GL.3).[1] Also, the ones

---

[1] Incidentally, it is possible to formulate hypersequent calculi for S4.2Grz, S4.3Grz,

for which we can show cut-elimination are the ones that have superintuitionistic counterparts via Gödel embedding. Hence, although this may be technically unsurprising, this can be taken as showing a conceptual stability of this variety of logics as "logic" compared with other modal logics. Semantically, the difference between these superintuitionistic logics (logic of weak excluded middle, Gödel-Dummett logic, classical logic) and our strict implication logics (SIS4.3, SIS5) is that the latter lacks truth persistence.[2] From this point of view, these logics may be worth being called sub-Gödel-Dummett logic or "sub-classical logic" (in the sense similar to subintuitionistic logic [137].) These phenomena must be, certainly, more than mere coincidence. There must be some conceptual reason for these, although we should stop speculating too much. We will leave more systematic studies in this direction to future research.

.1 Miscellaneous proof systems

In this appendix, we present some important logical systems or proof-frameworks that are mentioned in the main text of Chapter 2, which are not necessarily commonly known materials among logicians.

S4.2.1, S4.3.1. However, proving cut-elimination seems to be a challenging problem.

[2] Axiomatically, the difference is presence or absence of an axiom schema of the form $A \to (B \to A)$).

.1.1 Kosta Došen's higher-level sequent calculi

1. Structural rules

(Here $+$ means the disjoint union. $\cup$ means the union. Γ, Δ are called set terms, which stand for *sets* of formulas. Hence, we do not need contraction.)

$$\text{Ascending (A) } \frac{A^n}{\emptyset \vdash^{n+1} \{A^n\}} \text{ Descending (D) } \frac{\emptyset \vdash^{n+1} \{A^n\}}{A^n} \text{ Iteration (I) } \frac{A^n}{A^n}$$

$(n \geq 0)$

$$\text{Cut (C) } \frac{\Gamma \vdash^{n+1} \Delta + \{A^n\} \qquad \Theta + \{A^n\} \vdash^{n+1} \Xi}{\Gamma \cup \Theta \vdash^{n+1} \Delta \cup \Xi}, \; n \geq 0,$$

provided either $\Gamma \neq \Theta \cup \{A^n\}$ or $\Delta + \{A^n\} \neq \Xi$.

$$\text{Thinning (T) } \frac{\Gamma \vdash^{n+1} \Delta}{\Gamma \cup \Theta \vdash^{n+1} \Delta \cup \Xi}, \text{ provided either } \Gamma \neq \Gamma \cup \Theta \text{ or } \Delta \neq \Delta \cup \Xi.$$

Restrictions on thinning:

$$\text{Thinning}_H \text{ (for intuitionistic logic)} : \begin{cases} \text{if } \Delta = \emptyset, \Xi \text{ must be a singleton} \\ \quad \text{or empty;} \\ \text{if } \Delta \neq \emptyset, \Xi \text{ must be empty} \end{cases}$$

Thinning_K (for minimal logic) : Ξ must be empty.

Horizontalizing of rules: All the horizontalization of the level-preserving rule are axioms or axiom-schemata.

2. Rules for (classical) logical constants

$$(\to)\,\begin{array}{c}\Gamma + \{A\} \vdash^1 \Delta + \{B\}\\ \hline\hline \Gamma \vdash^1 \Delta + \{A \to B\}\end{array}$$

$$(\wedge)\,\begin{array}{c}\Gamma \vdash^1 \Delta + \{A\} \qquad \Gamma \vdash^1 \Delta + \{B\}\\ \hline\hline \Gamma \vdash^1 \Delta + \{A \to B\}\end{array}$$

$$(\wedge)\,\begin{array}{c}\Gamma \vdash^1 \Delta + \{A\} \qquad \Gamma \vdash^1 \Delta + \{B\}\\ \hline\hline \Gamma \vdash^1 \Delta + \{A \to B\}\end{array}$$

$$(\wedge)\,\begin{array}{c}\Gamma + \{A\} \vdash^1 \Delta \qquad \Gamma + \{B\} \vdash^1 \Delta\\ \hline\hline \Gamma \vdash^1 \Delta + \{A \to B\}\end{array}$$

$$(\bot)\,\begin{array}{c}\Gamma \vdash^1 \emptyset\\ \hline\hline \Gamma \vdash^1 \{\bot\}\end{array}$$

Intuitionistic (minimal) logic can be obtained as structural variants of classical logic by restricting Thinning as noted above.

3. Rules for strict implication and modal operator

$$(\Longrightarrow)\,\begin{array}{c}\Pi + \{A \vdash^1 B\} \vdash^2 \Sigma + \{\Theta \vdash^1 \Xi\}\\ \hline\hline \Pi \vdash^2 \Sigma + \{\Theta + \{A \Longrightarrow B\} \vdash^1 \Xi\}\end{array}$$

$$(\Box)\,\begin{array}{c}\Pi + \{\emptyset \vdash^1 A\} \vdash^2 \Sigma + \{\Theta \vdash^1 \Xi\}\\ \hline\hline \Pi \vdash^2 \Sigma + \{\Theta + \{\Box A\} \vdash^1 \Xi\}\end{array}$$

It is easy to observe that $\Box$ can be defined by $\Longrightarrow$ (with $\top$ in the language) (and vice versa).

The system with the modal rule with unrestricted thinning in level 2 would be a higher-level sequent calculus for $\mathsf{S5}$. But with restricted thinning in level 2, the system corresponds to $\mathsf{S4}$ modal logic.

Note that unlike classical or intuitionistic implication, the double-line rule for strict implication requires a sequent of level two. The rule for the classical (intuitionsitic) implication corresponds to Avron's biconditional for strong internal (internal) implication with additional structural conditions. Hence, they require only rules of level one. (In Došen's case, he formulates weaker logics by restricting weakening (thinning), instead of restricting the succedent into at most one formula.)

Among the implications discussed in chapter 2, one can see that Scott's system is the weakest, and the order of the rest is as follows: Došen's $\mathsf{S4}$ strict implication; Avron's biconditional for internal implication (formulate as set of formulas)[3]; Došen's rule for classical implication.

Scott's strict implication is not strong enough for MP2, and $\mathsf{S4}$ strict implication is just strong enough for it, but the use of level 2 rule strongly suggests that there is an important difference between $\mathsf{S4}$ strict implication and intuitionistic implication. The $\downarrow$ direction of it corresponds to the deduction theorem. $\mathsf{S4}$ strict implication is not strong enough for the deduction theorem to hold. (An answer to the question of which form of a rule cor-

---

[3] This is the same as the single-conclusion version of Došen's rule.

responding to the deduction theorem in S4 strict implication logic can be obtained without using the higher-level sequents is given in 2.4.3.2.b.) Also, Scott's strict implication is not strong enough to derive the ↑ direction, either, since MP2 is a special case of it. On the other hand, Došen's rule for S4 strict implication can derive it. In the following, we report the derivation, and the derivation of a level 2 analogue of the deduction theorem, which is closest to a plausible formulation of the definitional equation using nested $\vdash$ by Sambin et al.

Proposition .1.1 *Došen's rule ($\Longrightarrow$) for strict implication can derive the following two principle by using structural rules in Došen's higher-level sequent calculi.*

1. $A, A \Longrightarrow B \vdash^1 B$
2. $\Gamma \vdash^2 \{A \vdash^1 B\}$ *iff* $\Gamma \vdash^2 \{\emptyset \vdash^1 A \Longrightarrow B\}$

Proof We show these propositions by using Došen's system for S4 strict implication.

A proof of 1 looks as follows.

$$\dfrac{\dfrac{\{A \vdash^1 B\} \vdash^2 \{A \vdash^1 B\}}{\emptyset \vdash^2 \{A, A \Longrightarrow B \vdash^1 B\}}\,(\Longrightarrow)\!\downarrow}{A, A \Longrightarrow B \vdash^1 B}\,\text{Descending}$$

A proof of 2 looks as follows. First, we show the left-right direction.

$$\dfrac{\Gamma \vdash^2 \{A \vdash^1 B\} \qquad \dfrac{\dfrac{A \Longrightarrow B \vdash^1 A \Longrightarrow B}{\emptyset \vdash^2 \{A \Longrightarrow B \vdash^1 A \Longrightarrow B\}}\,\text{Ascending}}{\{A \vdash^1 B\} \vdash^2 \{\emptyset \vdash^1 A \Longrightarrow B\}}\,(\Longrightarrow)\uparrow}{\Gamma \vdash^2 \{\emptyset \vdash^1 A \Longrightarrow B\}}\,\text{Cut}$$

Then we show the converse direction. Using the second line of the proof of the previous proposition, i.e. $\emptyset \vdash^2 \{A, A \Longrightarrow B \vdash^1 B\}$, we can apply the following cut.

$$\frac{\Gamma \vdash^2 \{\emptyset \vdash^1 A \Longrightarrow B\} \qquad \emptyset \vdash^2 \{A, A \Longrightarrow B \vdash^1 B\}}{\Gamma \vdash^2 \{A \vdash^1 B\}}.$$

This suffices to show that $\dfrac{\Gamma \vdash^2 \{\emptyset \vdash^1 A \Longrightarrow B\}}{\Gamma \vdash^2 \{A \vdash^1 B\}}$.

Note that there is a slight but interesting difference between the proposition 2 and Sambin et al.'s crude formulation of the nested $\vdash$. In the latter case, $\vdash$ corresponding to $\vdash^1$ entirely disappears. Intuitively, this must be a phenomenon which makes the deduction theorem possible. But recall that the deduction theorem in the most general formulation does not hold for S4 strict implication. In this sense, strict implications are not the same as implications in substructural logics. A framework like the higher-level sequents may be useful to uniformly formulate strict implication and substructural logics.

.1.2 Sambin et al.'s system B

Just for reference, we present the following sequent calculus B in Sambin et al's program of basic logic for reference. (In their notation, $\vdash$ is like our $\Rightarrow$.) Γ, Δ, etc. are sequences of formulas. Hence, the structural exchange is explicitly formulated. All other structural rules are dropped in this basic systems and restored when we formulate stronger logics, such as linear, quantum, intuitionistic, relevant, classical logics. We do not discuss in Chapter 2 in detail, but the methodological emphasis on symmetry naturally motivate Sambin et al. to have co-implication in the system.

Axiom: $A \vdash A$

Structural rules:

exchange L $\dfrac{\Gamma, \Sigma, \Pi, \Gamma' \vdash \Delta}{\Gamma, \Pi, \Sigma, \Gamma' \Rightarrow \Delta}$ exchange R $\dfrac{\Gamma \vdash \Delta, \Pi, \Sigma, \Delta'}{\Gamma \vdash \Delta, \Sigma, \Pi, \Delta'}$

Operational rules

Multiplicatives:

formation L$\otimes$ $\dfrac{B, A \vdash \Delta}{B \otimes A \vdash \Delta}$ R$\oplus$ $\dfrac{\Gamma \vdash A, B}{\Gamma \vdash A \oplus B}$

reflection L$\oplus$ $\dfrac{B \vdash \Delta_1 \qquad A \vdash \Delta_2}{B \oplus A \vdash \Delta_1, \Delta_2}$ R$\otimes$ $\dfrac{\Gamma_2 \vdash A \qquad \Gamma_1 \vdash B}{\Gamma_1, \Gamma_2 \vdash A \otimes B}$

formation L1 $\dfrac{\vdash \Delta}{1 \vdash \Delta}$ R0 $\dfrac{\Gamma \vdash}{\Gamma \vdash 0}$

reflection L0 $0 \vdash$ R0 $\vdash 1$

Additives:

$$\text{formation} \quad \text{L}\vee\ \frac{B \vdash \Delta \qquad A \vdash \Delta}{B \vee A \vdash \Delta} \qquad\qquad \text{R}\wedge\ \frac{\Gamma \vdash A \qquad \Gamma \vdash B}{\Gamma \vdash A \wedge B}$$

$$\text{reflection} \quad \text{L}\wedge\ \frac{B \vdash \Delta}{B \wedge A \vdash \Delta} \qquad \frac{A \vdash \Delta}{B \wedge A \vdash \Delta} \qquad \text{R}\wedge\ \frac{\Gamma \vdash A}{\Gamma \vdash A \wedge B} \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \wedge B}$$

Implications:

$$\text{formation} \quad \text{L}\leftarrow\ \frac{B \vdash A}{B \leftarrow A \vdash} \qquad\qquad \text{R}\rightarrow\ \frac{A \vdash B}{\vdash A \rightarrow B}$$

$$\text{reflection} \quad \text{L}\leftarrow\ \frac{\vdash B \qquad A \vdash \Delta}{B \leftarrow A \vdash \Delta} \qquad\qquad \text{R}\leftarrow\ \frac{\Gamma \vdash A \qquad B \vdash}{\Gamma \vdash A \leftarrow B}$$

$$\text{order} \quad \text{U}\leftarrow\ \frac{D \vdash C \qquad B \vdash A}{D \leftarrow A \vdash C \leftarrow B} \qquad\qquad \text{U}\leftarrow\ \frac{A \vdash B \qquad C \vdash D}{B \rightarrow C \vdash A \rightarrow D}$$

.1.3 Wansing's display calculi for normal modal logics

The object language is the same as the usual language of modal logic. The deductive-metalanguage of the display calculi is a little idiosyncratic. It consists of the following items **I** (identity), $\circ$ (structural connective like comma in traditional sequents), $*$ (structural connective corresponding to "negation"), $\bullet$(structural connective for modality). We follow the notational convention in display calculi. A,B,C are formulas, X,Y,Z, are structures. $\rightarrow$ here means $\Rightarrow$ for our sequent calculi. (Although it may be obvious, let us note that $\vdash$ plays the role of the horizontal line. Interestingly, they do not have any particular symbol for juxtaposing two premises in cut.) The following are Wansing's display calculi for normal modal logic in [172]. Although we present cut here, cut is eliminable from these systems.

Logical rules: (Id) $\vdash p \rightarrow p$ (Cut) $X \rightarrow A \quad A \rightarrow Y \vdash X \rightarrow Y$.

Basic structural rules:

(1) $X \circ Y \rightarrow Z \dashv\vdash X \rightarrow X \circ Y^* \dashv\vdash Y \rightarrow X^* \circ Z$

(2) $X \rightarrow Y \circ Z \dashv\vdash X \circ Z^* \rightarrow Y \dashv\vdash Y^* \circ X \rightarrow Z$

(3) $X \rightarrow Y \dashv\vdash Y^* \rightarrow X^* \dashv\vdash X \rightarrow Y^{**}$

(4) $X \rightarrow \bullet Y \dashv\vdash \bullet X \rightarrow Y$,

$\dashv\vdash$ abbreviates both ways of $\vdash$.

Operational rules:

$(\rightarrow 1) \vdash \mathbf{I} \rightarrow 1$ $(1 \rightarrow)\ \mathbf{I} \rightarrow X \vdash 1 \rightarrow X$

$(\rightarrow 0) \vdash X \rightarrow \mathbf{I}$ $(0 \rightarrow)\ 0 \rightarrow \mathbf{I}$

$(\rightarrow \neg)\ X \rightarrow A^* \vdash X \rightarrow \neg A$ $(\neg \rightarrow)\ A^* \rightarrow X \vdash \neg A \rightarrow X$

$(\rightarrow \wedge)\ X \rightarrow A \quad Y \rightarrow B \vdash X \circ Y \rightarrow A \wedge B$

$(\wedge \rightarrow)\ A \circ B \rightarrow X \vdash A \wedge B \rightarrow X$

$(\rightarrow \vee)\ X \rightarrow A \circ B \vdash X \rightarrow A \vee B$

$(\vee \to)\quad A \to X \qquad B \to Y \vdash A \vee B \to X \circ Y$

$(\to\supset)\ X \circ A \to B \vdash X \to A \supset B$

$(\supset\to)\quad X \to A \qquad B \to Y \vdash A \supset B \to X^* \circ Y$

$(\to \Box)\ \bullet X \to A \vdash X \to \Box A \qquad (\Box \to)\ A \to Y \vdash \Box A \to \bullet Y$

$(\to \Diamond)\ X \to A \vdash (\bullet(X^*))^* \to \Diamond A \qquad (\Diamond \to)\ (\bullet(A^*))^* \to Y \vdash \Diamond A \to Y$

The following additional structural rules are sufficient.

$(\mathbf{I}+)\ X \to Z \vdash \mathbf{I} \circ X \to Z \qquad X \to Z \vdash \mathbf{I} \circ X \to Z$

$(\mathbf{I}-)\ \mathbf{I} \circ X \to Z \vdash X \to Z \qquad \mathbf{I} \circ X \to Z \vdash X \to Z$

$(\mathbf{1}ex)\ \mathbf{I} \to X \vdash Z \to X$

$(ex\mathbf{0})\ X \to \mathbf{I} \vdash X \to Z$

$(A)\ X_1 \circ (X_2 \circ X_3) \to Z \dashv\vdash (X_1 \circ X_2) \circ X_3 \to Z$

$(P)\ X_1 \circ X_2 \to Z \dashv\vdash X_2 \circ X_1 \to Z$

$(C)\ X \circ X \to Z \dashv\vdash X \to Z$

$(M)\ X_1 \to Z \dashv\vdash X_1 \circ X_2 \to Z$

$X_1 \to Z \dashv\vdash X_2 \circ X_1 \to Z$

$(MN)\ \mathbf{I} \to X \vdash \bullet\mathbf{I} \to X$

For each of modal axioms, we have the corresponding structural rule as follows.

$(D)\ \bullet X \circ \bullet \to \mathbf{I}^* \vdash X \to Y^*$

$(T)\ X \rightarrow \bullet Y \vdash X \rightarrow Y$

$(4)\ X \rightarrow \bullet Y \vdash X \rightarrow \bullet \bullet Y$

$(5)\ (\bullet(X^*))^* \rightarrow Y \vdash \bullet((\bullet(X^*))^*) \rightarrow Y.$

$(B)\ (\bullet(X^*))^* \rightarrow Y \vdash \bullet X \rightarrow Y$

.2 Cut-admissibility for a sequent calculus for SIS4 with $\supset$

It is convenient to semantically prove cut-admissibility for a traditional sequent calculus for S4 strict implication logic with material implication for two reasons. First, some derivations in the Hilbert-style axiomatic systems for strict implications are very tedious, so we want to appeal to sequent derivations as a proof for SIS4 case. Second, we discussed the issue of conservativeness in Chapter 2. We stated that adding classical implication to SIS4 logic is a conservative extension. Proving cut-elimination guarantees this.

Let $\mathsf{GSIS4}_{\supset}$ be a Gentzen-style sequent calculus for S4 strict implication with material implication given as follows. Here we use $\longrightarrow$ for strict implication and $\supset$ for material implication.

Rules for strict implication:

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

Rules for material implication:

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

We assume the following.

1. A sequent is a set, so the contraction rule is absorbed; we assume weakening.

2. Axioms and rules for $\bot$, $\top$, $\wedge$, $\vee$ are the same as presented above.

Semantic soundness and completeness will be proven with respect to the following Kripke semantics. Let be (K, R) a Kripke frame. K is a non-empty set, and R is a reflexive and transitive relation on K. Let $(K, R, \mathcal{V})$ be a Kripke model based on the Kripke frame (K, R) where $\mathcal{V}$ is a valuation on propositional variable to the subset of W. We define the relation $\Vdash$ satisfying the following properties. For any $u \in K$,

1. $u \Vdash p$ iff $u \in \mathcal{V}(p)$; $u \nVdash \bot$;

2. $u \Vdash \varphi \supset \psi$ iff, if $u \Vdash \varphi$, then $u \Vdash \psi$ (the standard clauses for $\wedge$, $\vee$);

3. $u \Vdash \varphi \longrightarrow \psi$ iff for any $v \in K$, s.t. uRv, if $v \Vdash \varphi$, then $v \Vdash \psi$.

For semantic soundness, we just state the proposition. (The validity of a sequent $\Gamma \Rightarrow \Delta$ is naturally defined from its translation to $\bigwedge \Gamma \Rightarrow \bigvee \Delta$.)

Proposition .2.1 (Soundness) *If* $\mathsf{GSIS4}_{\supset} \vdash \Gamma \Rightarrow \Delta$, *then* $\Gamma \Rightarrow \Delta$ *is valid.*

Now, we prove completeness for a cut-free sequent calculus GSIS4$_{\supset}^{-}$. We first define the notion of saturated sequents.

Definition .2.2 *1. If $A \longrightarrow B \in \Gamma$, then $B \in \Gamma$ or $A \in \Delta$;*

2. If $A \supset B \in \Gamma$, then $B \in \Gamma$ or $A \in \Delta$;

3. If $A \wedge B \in \Gamma$, then $A \in \Gamma$ and $B \in \Gamma$;

4. If $A \wedge B \in \Delta$, then $A \in \Delta$ or $B \in \Delta$;

5. If $A \vee B \in \Gamma$, then $A \in \Gamma$ or $B \in \Gamma$;

6. If $A \vee B \in \Delta$, then $A \in \Delta$ and $B \in \Delta$;

7. If $A \supset B \in \Delta$, then $A \in \Gamma$ and $B \in \Delta$.

The saturation algorithm is given as follows:

If $A \longrightarrow B \in \Gamma$, then put B to Γ or A to Δ;

If $A \supset B \in \Gamma$, then put B to Γ or A to Δ; If $A \supset B \in \Delta$, then A to Γ and B to Δ;

If $A \wedge B \in \Gamma$, then put A to Γ and B to Γ; If $A \wedge B \in \Delta$, then put A to Δ or B to Δ;

If $A \vee B \in \Gamma$, then put A to Γ or B to Γ; If $A \vee B \in \Delta$, then put A to Δ and B to Δ;

(after taking each step, the relevant formula will be inactive.)

If $\vdash \Gamma \Rightarrow \Delta$, then backtrack;

If $\nvdash \Gamma \Rightarrow \Delta$ with all the formulas become inactive, then stop.

Obviously, after the algorithm terminating successfully, the sequent satisfies the saturation property. Now we prove the theorem.

Lemma .2.3 (Saturation lemma) *If* $\nvdash \Gamma \Rightarrow \Delta$*, then there exists a saturated sequent* $\nvdash \Gamma'_i \Rightarrow \Delta'_i$*, s.t.* $\nvdash \Gamma'_i \Rightarrow \Delta'_i$*.*

Proof Proof by contradiction. Suppose $\nvdash \Gamma \Rightarrow \Delta$, but, after the algorithm terminating, we do not have any way of backtracking and $\vdash \Gamma' \Rightarrow \Delta'$. Observe that all the steps in the saturation algorithm are inverted operations of rules in the sequent calculus. This implies that if the saturated sequent is provable, then we can construct a cut-free proof of the original sequent $\Gamma \Rightarrow \Delta$. But this contradicts the assumption. ⊠

Consider $Sub(\Gamma \cup \Delta)$ (the set of subformulas of formulas occurring in $\Gamma \cup \Delta$).

Also, consider the set of all saturated sequents $\Gamma' \Rightarrow \Delta'$ consisting of formulas from $Sub(\Gamma \cup \Delta)$, s.t. $\nvdash \Gamma' \Rightarrow \Delta'$. Call it W. Define a relation among sequents in W as follows. $(\Gamma'_i \Rightarrow \Delta'_1)R(\Gamma'_j \Rightarrow \Delta'_j)$ iff $\{A \longrightarrow B | A \longrightarrow B \in \Gamma'_i\} \subseteq \Gamma'_j$. Then let the ordered pair (W, R) be our Kripke frame for SIS4. R is obviously reflexive and transitive. We define the forcing relation for propositional variables to be $(\Gamma'_i \Rightarrow \Delta'_i) \Vdash p$ iff $p \in \Gamma'_i$. $((\Gamma'_i \Rightarrow \Delta'_i \Vdash \bot$ iff $\bot \in \Gamma'_i)$.[4] Now we prove the semantic lemma.

Lemma .2.4 (semantic lemma) *1.* $\varphi \in \Gamma'_i \Longrightarrow (\Gamma'_i \Rightarrow \Delta'_1) \Vdash \varphi$*;*

2. $\varphi \in \Delta'_i \Longrightarrow (\Gamma'_i \Rightarrow \Delta'_1) \nVdash \varphi$*.*

---

[4] $\bot$ cannot occur any Γ'_i in any unprovable saturated sequent by construction.

Proof Proof by induction on the structure of formulas. Atomic cases are immediate by definition. We show only the case of strict implication.

Case 1. Suppose $\varphi_1 \longrightarrow \varphi_2 \in \Gamma'_i$ and suppose $(\Gamma'_i \Rightarrow \Delta'_i)R(\Gamma'_j \Rightarrow \Delta'_j)$. By definition, the latter implies $\varphi_1 \longrightarrow \varphi_2 \in \Gamma'_j$. By the saturation condition, this implies $\varphi_2 \in \Gamma'_j$ or $\varphi_1 \in \Delta'_j$. By IH, $(\Gamma'_j \Rightarrow \Delta'_j) \not\Vdash \varphi_1$ or $(\Gamma'_j \Rightarrow \Delta'_j) \Vdash \varphi_2$. Hence, $(\Gamma'_i \Rightarrow \Delta'_i) \Vdash \varphi_1 \longrightarrow \varphi_2$.

Case 2. Suppose $\varphi_1 \longrightarrow \varphi_2 \in \Delta'_i$. Let $\Gamma_k = \{A \longrightarrow B | A \longrightarrow B \in \Gamma'_i\} \cup \{\varphi_1\}$ and $\Delta_k = \{\varphi_2\}$. We claim that $\mathsf{SIS4}^-_\supset \nvdash \Gamma_k \Rightarrow \Delta_k$.

Suppose not. Then it would be that for all $A_1 \longrightarrow B_1, \ldots, A_n \longrightarrow B_n \in \Gamma_k$, $\vdash A_1 \longrightarrow B_1, \ldots, A_n \longrightarrow B_n, \varphi_1 \Rightarrow \varphi_2$. Then we can construct the following proof.

$$\dfrac{\dfrac{A_1 \longrightarrow B_1, \ldots, A_n \longrightarrow B_n, \varphi_1 \Rightarrow \varphi_2}{A_1 \longrightarrow B_1, \ldots, A_n \longrightarrow B_n \Rightarrow \varphi_1 \longrightarrow \varphi_2}}{\Gamma'_i \Rightarrow \Delta'_i}\ \text{weakening}$$

This contradicts the construction of the model, namely taking the unprovable saturated sequents. (Hence, $\nvdash \Gamma'_i \Rightarrow \Delta'_i$.) By the saturation lemma, $\Gamma_k \Rightarrow \Delta_k$ can be extended to an unprovable saturated sequent $\Gamma'_k \Rightarrow \Delta'_k$. Note that by construction $\{A \longrightarrow B | A \longrightarrow B \in \Gamma'_i\} \subseteq \Gamma'_k$. Hence $(\Gamma'_i \Rightarrow \Delta'_i)R(\Gamma'_k \Rightarrow \Delta'_k)$. Also, $\varphi_1 \in \Gamma'_k$ and $\varphi_2 \in \Delta'_k$. Hence, by IH, $(\Gamma'_k \Rightarrow \Delta'_k) \Vdash \varphi_1$ and $(\Gamma'_k \Rightarrow \Delta'_k) \not\Vdash \varphi_2$.

Thus, there exists $(\Gamma'_k \Rightarrow \Delta'_k) \in W$, s.t. $(\Gamma'_i \Rightarrow \Delta'_i)R(\Gamma'_k \Rightarrow \Delta'_k)$, $(\Gamma'_k \Rightarrow \Delta'_k) \Vdash \varphi_1$ and $(\Gamma'_k \Rightarrow \Delta'_k) \not\Vdash \varphi_2$. This means that $(\Gamma'_k \Rightarrow \Delta'_k) \Vdash \varphi_1 \longrightarrow \varphi_2$.

Theorem .2.5 (Completeness) [5] *If* $\Gamma \Rightarrow \Delta$ *is valid, then* $\mathsf{SIS4}^-_\supset \vdash \Gamma \Rightarrow \Delta$.

---

[5] By using this sequent calculus, we can easily prove the following in $\mathsf{SIS4}$. $A \longleftrightarrow B$ is

Proof Immediate consequences of the saturation lemma and the semantic lemma.

Corollary .2.6 (Conservativeness) $\mathsf{SIS4}_{\supset}$ *is a conservative extension of* $\mathsf{SIS4}$*.*

---

an abbreviation of $(A \longrightarrow B) \wedge (B \longrightarrow A)$.

1. $\mathsf{SIS4} \vdash ((A \longrightarrow B) \longrightarrow ((C \longrightarrow D) \longrightarrow E)) \longrightarrow ((C \longrightarrow D) \longrightarrow ((A \longrightarrow B) \longrightarrow E))$
2. $\mathsf{SIS4} \vdash (A \longrightarrow ((C \longrightarrow D) \longrightarrow E)) \longrightarrow ((C \longrightarrow D) \longrightarrow (A \longrightarrow E))$
3. $\mathsf{SIS4} \vdash ((A \longrightarrow B) \longrightarrow ((C \longrightarrow D) \longrightarrow E)) \longleftrightarrow ((A \longrightarrow B) \wedge (C \longrightarrow D) \longrightarrow E)$
4. $\mathsf{SIS4} \vdash (\bigwedge \Gamma_{\longrightarrow} \longrightarrow ((C \longrightarrow D) \longrightarrow E)) \longleftrightarrow (\bigwedge \Gamma_{\longrightarrow} \wedge (C \longrightarrow D) \longrightarrow E)$
5. $\mathsf{SIS4} \vdash (A \longrightarrow B) \longrightarrow ((B \longrightarrow C) \longrightarrow (A \longrightarrow C))$
6. $\mathsf{SIS4} \vdash ((((A \longrightarrow B) \longrightarrow \bot) \longrightarrow \bot) \longrightarrow \bot) \longrightarrow ((A \longrightarrow B) \longrightarrow \bot)$
7. $\mathsf{SIS4} \vdash (A \longrightarrow (A \longrightarrow B)) \longrightarrow (A \longrightarrow B)$

Bibliography

[1] A. R. Anderson and N. D. Belnap. *Entailment: The Logic of Relevance and Necessity*, volume 1. Princeton University Press, 1975.

[2] A. R. Anderson, N. D. Belnap, and J. M. Dunn. *Entailment: the Logic of Relevance and Necessity Vol 2.* Princeton University Press, 1992.

[3] S. Artemov and R. Iemhoff. The basic intuitionistic logic of proofs. *Journal of Symbolic Logic*, 72, 2007.

[4] S. N. Artëmov. Logic of proofs. *Ann. Pure Appl. Logic*, 67(1-3):29–59, 1994.

[5] S. N. Artemov. Deep isomorphism of modal derivations and lambda-terms, July 24 1999.

[6] S. N. Artemov. Explicit provability and constructive semantics. *The Bulletin of Symbolic Logic*, 7(1):1–36, 2001.

[7] S. N. Artemov and E. Nogina. Logic of knowledge with justifications from the provability perspective. Technical report, CUNY Ph.D. Program in Computer Science Technical Report TR-2004011, 2004.

[8] S. N. Artemov and E. Nogina. Introducing justification into epistemic logic. *J. Log. Comput*, 15(6), 2005.

[9] A. Avron. Tonk- a full mathematical solution.

[10] A. Avron. A constructive analysis of RM. *Journal of Symbolic Logic*, 52(4):939–951, 1987.

[11] A. Avron. Gentzenizing schroeder-heister's natural extension of natural deduction. *Notre Dame Journal of Formal Logic*, 31(1):127–135, 1990.

[12] A. Avron. Relevance and paraconsistency - a new approach. *J. Symb. Log.*, 55(2):707–732, 1990.

[13] A. Avron. Relevance and paraconsistency - a new approach, part ii: The formal systems. *Notre Dame Journal of Formal Logic*, 31(2):169–202, 1990.

[14] A. Avron. Hypersequents, logical consequence, and intermediate logics for concurrency. *Annals of Mathematics and Artificial Intelligence*, 4:225–248, 1991.

[15] A. Avron. Relevance and paraconsistency—A new approach. Part III: Cut-free Gentzen-type systems. *Notre Dame Journal of Formal Logic*, 32(1):147–160, Winter 1991.

[16] A. Avron. Simple consequence relations. *Information and Computation*, 92:105–139, 1991.

[17] A. Avron. Whither relevance logic? *Journal of Philosophical Logic*, 21:243–281, 1992.

[18] A. Avron. The method of hypersequents in the proof theory of propositional non-classical logics. In Wilfrid Hodges, Martin Hyland, Charles Steinhorn, and John Truss, editors, *Logic: from foundations to applications. Proc. Logic Colloquium, Keele, UK, 1993*, pages 1–32. Oxford University Press, New York, 1996.

[19] A. Avron. Multiplicative conjunction as an extensional conjunction. *Logic Journal of the IGPL*, 5(2), 1997.

[20] A. Avron. Negation: Two points of view. In D. Gabbay and H. Wansing, editors, *What is Negation?* Kluwer, 1999.

[21] R. Barcan Marcus. Strict implication, deducibility and the deduction theorem. *The Journal of Symbolic Logic*, 18(3):234–236, 1953.

[22] J. C. Beall and G. Restall. *Logical Pluralism.* Oxford Clarendon Press, 2006.

[23] N. D. Belnap. Tonk, plonk and plink. *Analysis*, 22(6):130–134, 1962.

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

[25] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic.* Cambridge University Press, 2001.

[26] E. Bonnay and B. Simmenauer. Tonk strikes back. *The Australasian Journal of Logic*, 3:33–44, 2005.

[27] G. Boolos. *The logic of provability.* Cambridge University Press, 1992.

[28] R. B. Brandom. *Making It Explicit: Reasoning, Representing, and Discursive Commitment.* Harvard University Press, 1994.

[29] K. Brünnler. Nested sequents (habilitationsschrift), 2010.

[30] J. P. Burgess. Which modal logic is the right one? *Notre Dame Journal of Formal Logic*, 40(1):81–93, 1999.

[31] R. Carnap. *Formalization of Logic.* Harvard University Press, Cambridge, Mass., 1943.

[32] A. Chagrov and M. Zakharyaschev. *Modal Logic.* Oxford University Press, Oxford, 1996.

[33] A. Ciabattoni. A proof-theoretical investigation of global intuitionistic (fuzzy) logic. *Archive for Mathematical Logic*, 44(4):435–457, 2004.

[34] A. Ciabattoni, N. Galatos, and K. Terui. From axioms to analytic rules in nonclassical logics. In *23th Symp. on Logic in Computer Science*, pages 229–240. IEEE Computer Society Press, 2008.

[35] A. Ciabattoni, G. Metcalfe, and F. Montagna. Algebraic and proof-theoretic characterizations of truth stressers for MTL and its extensions. *Fuzzy Sets and Systems*, 161(3):369–389, 2010.

[36] A. Ciabattoni, L. Straßburger, and K. Terui. Expanding the realm of systematic proof theory. In Erich Grädel and Reinhard Kahle, editors, *CSL*, volume 5771 of *Lecture Notes in Computer Science*, pages 163–178. Springer, 2009.

[37] G. Corsi. Weak logics with strict implication. *Zeitschr. f. math.Logik und Grundlagen d. Mathematik*, 33, 1987.

[38] H. B. Curry. The elimination theorem when modality is present. *J. Symb. Log*, 17(4):249–265, 1952.

[39] A. Dabrowski, L.S. Moss, and R. Parikh. Topological reasoning and the logic of knowledge. *Ann. Pure Appl. Logic*, 78(1-3):73–110, 1996.

[40] E. Dashkov. Arithmetical completeness of the intuitionistic logic of proofs. *Journal of Logic and Computation*, Advanced Access, August 2009.

[41] N. Dershowitz and D. A. Plaisted. Rewriting. In A. Robinson and A. Voronkov, editors, *Handbook of Automated Reasoning*, volume I, chapter 9, pages 535–610. Elsevier Science, 2001.

[42] K. Došen. *Logical Constants: An Essay in Proof Theory.* PhD thesis, University of Oxford, 1980.

[43] K. Došen. Minimal modal systems in which heyting and classical logic can be embedded. *Publications de l'Institut Mathmatique*, 44(30), 1981.

[44] K. Došen. Sequent-systems for modal logic. *J. Symb. Log*, 50(1):149–168, 1985.

[45] K. Došen. Logical constants as punctuation marks. *Notre Dame Journal of Formal Logic*, 30(3):362–381, 1989.

[46] K. Došen and P. Schroeder-Heister. Conservativeness and uniqueness. *Theoria*, 1985.

[47] K. Došen. On passing from singular to plural consequences.

[48] K. Došen. Sequent-systems and groupoid models. I. *Studia Logica: An International Journal for Symbolic Logic*, 47(4).

[49] K. Došen. Modal logic as metalogic. *Journal of Logic, Language, and Information*, 1(3):173–201, 1992.

[50] K. Došen. Modal translations in substructural logics. *Journal of Philosophical Logic*, 21(3):283–336, 1992.

[51] K. Došen. Logical consequence : a turn in style. In *Tenth Intern. Congress of Logic Methodology and Philosophy of Science (Dalla Chiara, M.L. and Doets, K. and Mundici, D. and van Benthem, J. eds.), Florence* , pages 289–311, Dordrecht, août 1997. Kluwer. (Volume 1) Dates de conférence : août 1997 1997.

[52] M. Dummett. The justification of deduction. *Proceedings of the British Academy*, 59.

[53] M. Dummett. *The Logical Basis of Metaphysics*. Duckworth, London, 1991.

[54] M. A. E. Dummett. *Frege: Philosophy of Language*. Duckworth, London, 1973.

[55] L. Fariñas del Cerro and A. Herzig. Combining classical and intuitionistic logic, or: Intuitionistic implication as a conditional. In F. Baader and K. U. Schulz, editors, *Frontiers of Combining Systems: Proceedings of the 1st International Workshop, Munich (Germany)*, pages 93–102. Kluwer Academic Publishers, 1996.

[56] M. Fitting. Logics with several modal operators. *Theoria*, 35.

[57] M. Fitting. *Proof Methods for Modal and Intuitionistic Logic*. Reidel Publishing Company, 1983.

[58] M. Fitting. Semantics and tableaus for LPS4. 2004.

[59] M. Fitting. Modal proof theory. In *Handbook of Modal Logic*. Elsevier, New York, 2006.

[60] C. Franks. Cut as consequence. 2009.

[61] G. Gentzen. *Mathematische Zeitschrift.*

[62] D. Gabbay. *Semantical Investigations in Heyting's Intuitionistic Logic.* Reidel, Dordrecht, Netherlands, 1981.

[63] D. M. Gabbay. *Labelled Deductive Systems*, volume 1—Foundations. Oxford University Press, 1996.

[64] N. Galatos, P. Jipsen, T. Kowalski, and H. Ono. *Residuated Lattices: An Algebraic Glimpse at Substructural Logics.* Elsevier, 2007.

[65] J. Garson. Natural semantics. *Theoria*, 67.

[66] S. Ghilardi. Incompleteness results in kripke semantics. *J. Symb. Log.*, 56(2):517–538, 1991.

[67] J. Girard. Linear logic. *Theoretical Computer Science*, 50:1–102, 1987.

[68] J.-Y. Girard. A new constructive logic: classical logic. *Mathematical Structures in Computer Science*, 1(3):255–296, November 1991.

[69] J.-Y. Girard. From foundations to ludics. *Bulletin of Symbolic Logic*, 79(2), 2003.

[70] K. Gödel. Vortrag bei Zilsel (1938). In S. Feferman et al., editor, *Kurt Gödel, Collected Works, vol. III.* Oxford Univ. Press, New York, 1995.

[71] M. Gomez-Torrente. The problem of logical constants. *Bulletin of Symbolic Logic*, 8(1):1–37, 2002.

[72] N. D. Goodman. A theory of constructions equivalent to arithmetic. In A. Kino, J. Myhill, and R. E. Vesley, editors, *Proceedings Summer Conf. on Intuitionism and Proof Theory, Buffalo, NY, USA, Aug 1968*, Studies in Logic and the Foundations of Mathematics, pages 101–120. North-Holland, Amsterdam, 1970.

[73] V. Goranko and M. Otto. Model theory of modal logic. In F. Wolter (eds.) P. Blackburn, J. van Benthem, editor, *Handbook of Modal Logic*, pages 249–329. Kluwer, 2007.

[74] V. Goranko and S. Passy. Using the universal modality: Gain and questions. *Journal of Logic and Computation*, 2 (1):5–30, 1992.

[75] R. Goré. Tableau methods for modal and temporal logics. In R Haehnle J Posegga (Eds.) M D'Agostino, D Gabbay, editor, *Handbook of Tableau Methods*, pages 297–396. Kluwer, 1999.

[76] R. Goré. Displaying modal logic, heinrich wansing. *Journal of Logic, Language and Information*, 9(2):269–272, 2000.

[77] R. Goré and R. Ramanayake. Valentini's cut-elimination for provability logic resolved. In Carlos Areces and Robert Goldblatt, editors, *Advances in Modal Logic*, pages 67–86. College Publications, 2008.

[78] D. Guaspari and R. M. Solovay. Rosser sentences. *Ann. Math. Logic*, 16(1):81–99, 1979.

[79] E. Gullberg and S. Lindström. Semantics and the justification of deductive inference. In J. Josefsson T. Rønnow-Rasmussen, B. Petersson and D. Egonsson (eds.), editors, *Hommage è Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowicz.* www.fil.lu.se/hommageawlodek.

[80] I. Hacking. What is logic? *Journal of Philosophy*, 76(6):285–319, 1979.

[81] H. Hodes. On the sense and reference of a logical constant. *Philosophical Quarterly*, 54:134–165, 2004.

[82] G. Huet. Confluent reduction: abstract properties and applications to term rewrite systems. *Journal of the Association for Computing Machinery*, 27(4):787–821, October 1980.

[83] L. Humberstone. For want of an 'and': A puzzle about non-conservative extension. *History and Philosophy of Logic*, 76(6):285–319, 1979.

[84] L. Humberstone. On a conservative extension argument of Dana Scott. *Logic Journal of the IGPL*, 19(1):241–288, 2011.

[85] L. Humberstone. *The Connectives.* MIT Press, forthcoming.

[86] A. Indrzejczak. Labelled analytic tableaux for S4.3. *Bulletin of the Section of Logic*, 31(1):15–26, 2002.

[87] R. Ishigaki and R. Kashima. Sequent calculi for some strict implication logics. *Logic Journal of the IGPL*, 16(2):155–174, 2008.

[88] G. Jäger. Modal Fixed Point Logics. In J. Esparza, B. Spanfelner, and O. Grumberg, editors, *Logics and Languages for Reliability and Security*, volume 25 of *NATO Science for Peace and Security Series - D: Information and Communication Security*. IOS Press, IOS Press, 2010.

[89] E. Jeřábek. A note on Grzegorczyk's logic. *Mathematical Logic Quarterly*, 50(3):295–296, 2004.

[90] R. Kahle and P. Schroeder-Heister. Introduction: Proof-theoretic semantics. *Synthese*, 148(3):503–506, 2006.

[91] R. Kashima and N. Kamide. Substructural implicational logics including the relevant logic E. *Studia Logica*, 63(2):181–212, 1999.

[92] A. Koslow. The implicational nature of logic: A structuralist account. *European Review of Philosophy, The Nature of Logic (A. C. Varzi, Ed.)*, 1999.

[93] G. Kreisel. Foundations of intuitionistic logic. In E. Nagel, P. Suppes, and A. Tarski, editors, *LMPS*, pages 198–210. Stanford University Press, Stanford, 1962.

[94] H. Kremer. Logic and meaning: The philosophical significance of the sequent calculus. *Mind*, 98:50–72, 1988.

[95] S. A. Kripke. Semantic considerations on modal logic. *Acta Philosophica Fennica*, 24:83–94, 1963.

[96] N. V. Krupski. On the complexity of the reflected logic of proofs. *Theor. Comput. Sci.*, 357(1-3):136–142, 2006.

[97] H. Kurokawa. Hypersequent calculi for intuitionistic logic with classical atoms. accepted for publication by *Annals of Pure and Applied Logic*.

[98] H. Kurokawa. Tableaux and hypersequents for justification logic. In *Logical Foundations of Computer Science*, pages 318–331, 2009.

[99] R. Kuznets. Self-referential justifications in epistemic logic. *Theory of Computing Systems*, 46(4):636–661, may 2010.

[100] S. Maehara. Eine darstellung der intuitionistischen logik in der klassischen. *Nagoya Mathematical Journal*, pages 45–64, 1954.

[101] P. Martin-Löf. On the meanings of the logical constants and the justifications of the logical laws. *Nordic Journal of Philosophical Logic*, 1(1):11–60, 1996. Lecture notes to a short course at Università degli Studi di Siena, April 1983.

[102] F. Massacci. Single step tableaux for modal logics. *J. Autom. Reasoning*, 24(3), 2000.

[103] G. Metcalfe. Proof theory for casari's comparative logics. *J. Log. Comput.*, 16(4):405–422, 2006.

[104] G. Metcalfe, N. Olivetti, and D. M. Gabbay. *Proof Theory for Fuzzy Logics*. Springer, 2009.

[105] G. Mints. Indexed systems of sequents and cut-elimination. *Journal of Philosophical Logic*, 26(6):671–696, 1997.

[106] G. Mints. Cut elimination for S4C: A case study. *Studia Logica*, 82(1):121–132, 2006.

[107] N. Negri. Proof analysis in modal logic. *Journal of Philosophical Logic*, 34(5-6):507–544, 2005.

[108] S. Negri and J. von Plato. *Structural Proof Theory.* Cambridge University Press, Cambridge, UK, 2001.

[109] E. Nogina. Epistemic completeness of GLA. *The Bulletin of Symbolic Logic*, 13(3):407, 2007.

[110] E. Nogina. Logic of strong provability and explicit proofs. *Proceedings of Logic Colloquium 2008*, 2008.

[111] P. W. O'Hearn and D. J. Pym. The logic of bunched implications. *Bulletin of Symbolic Logic*, 5(2):215–244, 1999.

[112] M. Okada. A weak intuitionistic propositional logic with purely constructive implication. *Studia Logica*, 46(4), 1987.

[113] M. Okada. On a theory of weak implications. *J. Symb. Log*, 53(1):200–211, 1988.

[114] M. Okada. A uniform semantic proof for cut-elimination and completeness of various first and higher order logics. *Theor. Comput. Sci.*, 281(1-2):471–498, 2002.

[115] K. Onishi, M. Matsumoto. Gentzen method in modal calculi i. *Osaka Mathematical Journal*, 1957.

[116] K. Onishi, M. Matsumoto. Gentzen method in modal calculi ii. *Osaka Mathematical Journal*, 1959.

[117] K. Onishi, M. Matsumoto. A system for strict implication. *Annals of the Japanese Association for Philosophy of Science*, 1964.

[118] F. Paoli. *Substructural Logics: A Primer*, volume 13 of *Trends in Logic*. Kluwer Academic Pub., 2002.

[119] F. Paoli. Quine and Slater on paraconsistency and deviance. *Journal of Philosophical Logic*, 32(2):531–548, 2003.

[120] F. Paoli. Implicational paradoxes and the meaning of logical constants. *Australasian Journal of Philosophy*, 85(4):553 – 579, 2007.

[121] R. Parikh. The completeness of propositional dynamic logic. In Józef Winkowski, editor, *MFCS*, volume 64 of *Lecture Notes in Computer Science*, pages 403–415. Springer, 1978.

[122] J. Peregrin. What is the logic of inference? *Studia Logica*, 88(2):263 – 294, 2008.

[123] F. Poggiolesi. Sequent calculi for modal logic (ph.d thesis), 2008.

[124] D. Prawitz. *Natural Deduction.* Almqvist & Wiksell, Uppsala, 1965.

[125] D. Prawitz. Meanings and proofs: On the conflict between classical and intuitionistic logic. *Theoria*, 43(1):2–40, 1977.

[126] D. Prawitz. Remarks on some approaches to the concept of logical consequence. *Synthese*, 62, 1985.

[127] D. Prawitz. Meaning Theory and Anti-realism. In B. McGinness and G. Oliveri, editors, *The Philosophy of Michael Dummett*, pages 79–89. Kluwer Academic Publishers, Netherland, 1994.

[128] D. Prawitz. Review of Michael Dummett "The Logical Basis of Metaphysics". *Mind*, 103, 1994.

[129] D. Prawitz. Meaning Approached via Proofs. *Synthese*, 148:507–524, 2006.

[130] G. Priest. *An Introduction to Non-classical Logic.* Cambridge University Press, 2001.

[131] A. N. Prior. The runabout inference-ticket. *Analysis*, 21(2):38–39, December 1960.

[132] S. Read. *Relevant Logic: A Philosophical Examination of Inference*. B. Blackwell, 1988.

[133] S. Read. Harmony and Modality. In L. Kieff C. Dgremont and H. Rckert, editors, *Dialogues, Logics and Other Strong Things: Essays in Honour of Shahid Rahman*, pages 285–303. College Publications, 2008.

[134] S. Read. General-elimination harmony and the meaning of the logical constants. *Journal of Philosophical Logic*, 39(2):557–76, 2010.

[135] B. Renne. Semantic cut-elimination for two explicit modal logics. In Janneke Huitnink and Sophia Katrenko, editors, *Proceedings of the 11th ESSLLI Student Session*, pages 148–158, Málaga, Spain, 2006.

[136] B. Renne. Public communication in justification logic. Technical Report TR-2005025, CUNY Ph.D. Program in Computer Science, 2007.

[137] G. Restall. Subintuitionistic logics. *Notre Dame Journal of Formal Logic*, 35(1):116–129, 1994.

[138] G. Restall. *An Introduction to Substructural Logics*. Routledge, February 2000.

[139] G. Restall. Multiple Conclusions. In Luis Valdes-Villanueva Petr Hajek and Dag Westerstahl, editors, *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress*, pages 189–205. Kings College Publications, 2005.

[140] G. Restall. Proofnets for S5: sequents and circuits for modal logic. In L. Newelski C. Dimitracopoulos and D. Normann, editors, *Logic Colloquium 2005*, pages 151–172. Cambridge University Press, 2007.

[141] G. Restall. Proof Theory and Meaning: the context of deducibility. In P. Maddy F. Delon, U. Kohlenbach and F. Stephan, editors, *Logic Colloquium 2007*, pages 204–219. Cambridge University Press, 2010.

[142] G. Sambin. Two applications of dynamic constructivism: Brouwers continuity principle and choice sequences in formal topology. In Michel Bourdeau Mark van Atten, Pascal Boldini and Gerhard Heinzmann, editors, *One Hundred Years of Intuitionism (1907-2007) The Cerisy Conference*, pages 41–51. Springer, 2008.

[143] G. Sambin, G. Battilotti, and C. Faggian. Basic logic: Reflection, symmetry, visibility. *Journal of Symbolic Logic*, 65(3):979–1013, 2000.

[144] P. Schoeder-Heister. Sequent calculi and bidirectional natural deduction: On the proper basis of proof-theoretic semantics. In *The Logica Yearbook 2008.* College Publications, 2009.

[145] P. Schroeder-Heister. Definitional reflection and basic logic. *To appear in the APAL.*

[146] P. Schroeder-Heister. A natural extension of natural deduction. *J. Symb. Log.*, 49(4):1284–1300, 1984.

[147] P. Schroeder-Heister. Structural frameworks with higher-level rules. philosophical investigations on the foundations of formal reasoning. (habilitationsschrift), 1987.

[148] D. Scott. On engendering an illusion of understanding. *Journal of Philosophy*, 68:787–807, 1971.

[149] D. Scott. Background to formalization. In H. Leblanc, editor, *Truth, Syntax and Modality: Proceedings of the Temple University Conference on Alternative Semantics*, pages 244–273. North-Holland, 1973.

[150] D. Scott. Completeness and axiomatizability in many-valued logic. In Leon Henkin *et al.*, editor, *Proceedings, Tarski Symposium*, pages 411–435. American Mathematical Society, 1974.

[151] D. Scott. Rules and derived rules. In S. (eds.) Stenland, editor, *Logical Theory and Semantic Analysis*, pages 147–161. Reidel, 1974.

[152] T. Shimura. Cut-free systems for the modal logic S4.3 and S4.3Grz. *Reports on Mathematical Logic*, 25, 1991.

[153] D. J. Shoesmith and T. J. Smiley. *Multiple-Conclusion Logic*. Cambridge University Press, 1978.

[154] J. Slaney. A general logic. *Australasian Journal of Philosophy*, 68(1):74 – 88, 1990.

[155] T. Smiley. The independence of connectives. *J. Symb. Log*, 27(4):426–436, 1962.

[156] O. Sonobe. A Gentzen-type formulation of some intermediate propositional logics. *J. Tsuda College*, 7, 1975.

[157] M. H. Sørensen and P. Urzyczyn. *Lectures on the Curry-Howard Isomorphism.* Elsevier, 2006.

[158] F. Steinberger. Tennant on multiple conclusions. *Logique et analyse*, 51(201), 2008.

[159] W. R. Stirton. Some problems for proof-theoretic semantics. *The Philosophical Quarterly*, 58(231):278–298, 2008.

[160] G. Sundholm. Hacking's logic. *The Journal of Philosophy*, 78(3):160–168, 1981.

[161] M.E. Szabo, editor. *The collected papers of Gerhard Gentzen.* North-Holland Pub. Co., Amsterdam, 1969.

[162] W. Tait. Gödel's interpretation of intuitionism. *Philosophia Mathematica*, 14(2):208–228, 2006.

[163] A. Tarski. Fundamental concepts of the methodology of the deductive sciences. In Alfred Tarski, editor, *Logic, Semantics, Metamathematics*, pages 60–109. Oxford University Press, Oxford, 1956. Translated by J.H. Woodger.

[164] A. Tarski. On the concept of logical consequence. In Alfred Tarski, editor, *Logic, Semantics, Metamathematics*, pages 409–420. Oxford University Press, Oxford, 1956. Translated by J.H. Woodger.

[165] N. Tennant. *Anti-Realism and Logic: Truth as Eternal.* Oxford University Press, Oxford, 1987.

[166] N. Tennant. *The Taming of the True.* Oxford University Press, Oxford, 1997.

[167] A. Troelstra and H. Schwichtenberg. *Basic Proof Theory.* Cambridge University Press, second edition, 2000.

[168] A. Troelstra and D. van Dalen. *Constructivism in Mathematics*, volume I, II. North Holland, 1988.

[169] A.M. Ungar. *Normalization, Cut-eliminations and the Theory of Proofs.* CSLI Lecture Notes, 1992.

[170] S. Valentini. A syntactic proof of cut-elimination for gl_{lin}. *Zeitschrift fr Mathematische Logik und Grundlagen der Mathematik*, 32, 1986.

[171] J. von Plato. Natural deduction with general elimination rules. *Arch. Math. Log.*, 40(7):541–567, 2001.

[172] H. Wansing. Sequent calculi for normal modal proposisional logics. *Journal of Logic and Computation*, 4(2):125–142, April 1994.

[173] H. Wansing. *Displaying Modal Logic.* Kluwer Academic Publishers, Dordrecht, 1998.

[174] H. Wansing. Translation of hypersequents into display sequents. *Logic Journal of the IGPL*, 6(5):719–733, 1998.

[175] H. Wansing. The idea of a proof-theoretic semantics and the meaning of the logical operations. *Studia Logica*, 64(1):3–20, 2000.

[176] S. Weinstein. The intended interpretation of intuitionistic logic. *Journal of Philosophical Logic*, 12(2):261–270, 1983.

[177] F. Wolter and M. Zakharyaschev. Intuitionistic modal logic. In P. Minari, editor, *Proceedings of the 10th International Congress of Logic, Methodology and Philosophy of Science.* Kluwer Academic Publishers, 1995.

[178] T. Yavorskaya. Logic of proofs and provability. *Ann. Pure Appl. Logic*, 113(1-3), 2001.

[179] J. Zucker. The correspondence between cut-elimination and normalisation. *Annals of Mathematical Logic*, 7:1–112, 1974.