Logical Methods in Computer Science
Volume 16, Issue 3, 2020, pp. 10:1–10:66
https://lmcs.episciences.org/

Submitted Aug. 07, 2018
Published Aug. 19, 2020

DUAL-CONTEXT CALCULI FOR MODAL LOGIC

G. A. KAVVOS

Institut for Datalogi, Aarhus University, Åbogade 34, 8200 Aarhus N, Denmark
e-mail address: alex.kavvos@cs.au.dk

Abstract. We present natural deduction systems and associated modal lambda calculi for the necessity fragments of the normal modal logics K, T, K4, GL and S4. These systems are in the dual-context style: they feature two distinct zones of assumptions, one of which can be thought as modal, and the other as intuitionistic. We show that these calculi have their roots in in sequent calculi. We then investigate their metatheory, equip them with a confluent and strongly normalizing notion of reduction, and show that they coincide with the usual Hilbert systems up to provability. Finally, we investigate a categorical semantics which interprets the modality as a product-preserving functor.

Introduction

The developments that have taken place over the past twenty years have shown that *constructive modalities*—broadly construed as unary type operators—are an important and versatile tool for both type theory and programming language theory.

Modalities have been used for various purposes within dependent type theory, e.g. to express proof irrelevance [Pfe01, AB04] and truncation in Homotopy Type Theory [RSS20], to formalise Cartan geometry [Wel17] and quantum gauge field theory [SS14], to internalise parametricity arguments in dependent type theory [NVD17, ND18], to reason about differential cohesive toposes [GLN$^+$17], to formally prove theorems that relate topology and homotopy [Shu18], and to construct models of universes internal to a topos [LOPS18].

To name but a few occurrences in programming language theory, modalities have been used in staged metaprogramming [DP01, TI10, Dav17], to control the complexity of typed programs [Hof99b], to enable recursion over higher-order abstract syntax [SDP01] and to provide categorical models for it [Hof99a], to control information flow [ABHR99, SI08, Kav19], to design a λ-calculus for distributed computing and mobile code [MCHP04], to build models of stateful languages with recursive types [BMSS12], in coinductive programming and guarded recursion [AM13, CBBB16], in functional reactive programming [KB11b, KB11a, Kri13, BGM19], in modelling contextual computation [Orc14], and, finally, in compartmentalising effects [Mog91] and combining them with resources [CFMM16].

Key words and phrases: modality, modal logic, modal type theory, Curry Howard correspondence, dual context, natural deduction, proof theory, categorical semantics, product-preserving functor, comonads.

This is a revised and extended version of a paper presented at LICS 2017 [Kav17].

LOGICAL METHODS IN COMPUTER SCIENCE
DOI:10.23638/LMCS-16(3:10)2020
© G. A. Kavvos
Creative Commons

Despite this wide applicability, there have been very few foundational studies on the *Curry-Howard-Lambek correspondence* that underlies modal types: we have surveyed relevant work in [Kav16]. The major impediment to carrying out such work is that the methods of modal proof theory are at best kaleidoscopic: while one type of calculus may work well for a specific logic, it may readily fail to express a closely related one. It is possible to develop a working intuition for these patterns, but it is much harder to precisely explain the root of these difficulties.

This paper has two goals. The first is to explain why a particular pattern of natural deduction for modal logics—namely that of *dual contexts*—is suited to the necessity fragments of the most popular *normal modal logics*. The key intuition is that the separation of assumptions into a modal zone and an intuitionistic zone allows us to mimic rules from cut-free sequent calculi for these logics. The second goal is to extend the Curry-Howard-Lambek correspondence and its usual triptych of logic, computation, and categories to normal modal logics. We show that the dual-context term languages admit a categorical semantics in which the modality is interpreted by a *product-preserving endofunctor* equipped with gadgets corresponding to the sundry axioms of normal modal logic. This deviates from previous approaches (in particular that of Bierman and de Paiva [BdP00]) but leads to a simpler syntax that nonetheless encompasses a large number of categorical models.

A road map. Our investigation is structured as follows. First, we define constructive versions of the most basic normal modal logics, namely K, T, $\mathsf{K4}$, GL and $\mathsf{S4}$, and present a Hilbert system for each (§1). We then briefly recount previous attempts at presenting natural deduction systems for them. This discussion leads us to a systematic method for deriving dual-context systems (§2).

Then, in §3 we reformulate these systems as modal λ-calculi, and study their basic metatheory. By writing down terms of the appropriate type, we show that our dual-context systems are equivalent to the Hilbert systems given in §2 at the level of provability. Following that, we endow these terms with a notion of reduction (§4). We prove that this has the usual good properties: it is confluent, strongly normalizing, and eliminates all cuts, i.e. the normal forms satisfy the subformula property. We stop short of deciding equality.

Finally, we give a categorical semantics for these calculi. We first introduce the relevant category theory in §5: this consists of a self-contained account of strong monoidal functors between cartesian categories, which we prove coincide with product-preserving functors. We then define various other gadgets used for interpreting the rules of our calculi, including coherent comultiplications (for $\mathsf{K4}$) and counits (for T). In the case of $\mathsf{S4}$, which contains both of these logics, these gadgets are required to satisfy the usual coherence equations of a product-preserving *comonad*. Finally, GL requires a novel notion of *modal fixed point*. We define a sound and complete interpretation into these structures in §6.

The case of Gödel and Löb. Perhaps one of the most interesting and surprising aspects of our investigation is that the general pattern described in §2 can be used to derive a natural deduction system for a constructive version of the *logic of provability*. While the classical version of the logic itself is comprehensively covered in the book by Boolos [Boo94], numerous interesting historical and mathematical facts about its constructive and intuitionistic variants are collected in a survey by Tadeusz Litak [Lit14].

GL is the least normal modal logic containing the *Gödel-Löb axiom*

$$\Box(\Box A \to A) \to \Box A \tag{$\star$}$$

for every formula A. If we erase the modal operators from this axiom we obtain the type $(A \to A) \to A$ of *fixed point combinators* at A. It is thus natural to suspect that, from a computational point of view, GL should lead to a calculus equipped with some unusual, modal kind of *recursion*.

This observation, which is originally due to Nakano [Nak00], led to a fruitful research programme on what is variously known as *productive coprogramming* [AM13], *guarded recursion* [BMSS12, BM13, CBBB16, Gua18], or *corecursion* [SdV12]. However, the underlying logic of these systems is essentially a variant of the *strong Gödel-Löb logic* SL, which is the least normal modal logic closed under the stronger axiom

$$(\Box A \to A) \to A$$

This axiom, which implies ($\star$) as well as the unusual formula $A \to \Box A$, is incompatible with classical normal modal logic, as no interesting Kripke frames satisfy the latter formula [Lit14, Remark 17]. Nevertheless, it appears time and again in various intutionistic settings.

Returning to the weaker logic GL, we find that our methods apply to yield a new natural deduction system, which—we argue in §2.1—is significantly simpler to the only other known system for GL, viz. that of Bellin [Bel85]. Given the intricate relationship between constructive provability logics, guarded recursion, and coinductively-defined infinite data, it may come as a surprise that the detour-eliminating reduction for this system is strongly normalizing. We will discuss that in more detail, but the basic intuition is that the introduction term in our calculus has a coinductive behaviour: it recursively unfolds only when forced to do so by an occurrence of the elimination rule.

Related work on dual contexts. Dual-context calculi were pioneed by Girard [Gir93], Andreoli [And92], Wadler [Wad93, Wad94], Plotkin [Plo93], and Barber [Bar96], in the setting of *multiplicative exponential linear logic* and its 'of course' (!) modality. They were then imported into the intuitionistic setting by Davies and Pfenning [PD01, DP01], who introduced the system for $\mathsf{S4}$ that we study in this paper. With the exception of $\mathsf{S4}$, the systems discussed in this paper are largely new. In some cases, glimpses of similar patterns have appeared before. Despite the popularity of the $\mathsf{S4}$ system in the programming language community, there has been no detailed study of its proof-theoretic properties.

An approach that is similar to ours for K and $\mathsf{K4}$ has been considered by Pfenning [Pfe13, Pfe15] in the context of *linear sequent calculi*, which seems to be closely related to the work of Danos and Joinet on *elementary linear logic* [DJ03]. However, this work of Pfenning remains unpublished, and the natural deduction system for K in this paper is independently due to the present author. The technical innovations needed in presenting a term calculus for $\mathsf{K4}$ and GL are new.

The study of reduction conducted by Davies and Pfenning for dual-context $\mathsf{S4}$ [PD01] was limited to an evaluation strategy used as operational semantics. In [DP01] a notion of conversion is introduced, but the study of its properties was left to a future paper, which never appeared. Neither of these papers discusses commuting conversions. The categorical interpretation of dual-context $\mathsf{S4}$ in terms of product-preserving comonads was briefly sketched by Hofmann [Hof99b, §2.6.12].

Related work on modal proof theory. The history of modal proof theory and constructive modal logics is long and tumultuous, so we shall try to avoid the subject as much as possible. A more thorough discussion of modal λ-calculi may be found in [Kav16]. For a broader survey of the proof theory of modal logic we recommend [Neg11].

While the earliest work goes back to Prawitz [Pra65], the first modal λ-calculus seems to be the Bierman-de Paiva system for $\mathsf{S4}$ [BdP92, BdP96, BdP00]. For reasons we discuss in §2, this system is unsatisfying: its proof theory requires many commuting conversions to eliminate all cuts, and its syntax is counterintuitive as a programming language. The proof-theoretic aspect is extensively discussed by Goubault-Larrecq [GL96, §5.2]. The programming-related issues are mentioned by Clouston et al. [CBBB16], who use a similar style of calculus in the context of guarded recursion. Clouston et al. try to argue that the "burden presented by the explicit substitutions seems quite small," but the fact this is a leitmotif in their paper significantly weakens their argument. Furthermore, in *op. cit.* it is argued that this style does not economically adapt to dependent types. Indeed, the dual-context type theory of Shulman [Shu18] seems to vindicate this claim.

There is little previous work on natural deduction for sub-$\mathsf{S4}$ systems of normal modal logic. Martini and Masini [MM96] presented a *Fitch-style* system. This was later adapted by Davies and Pfenning [DP01, §5], who call it the 'Kripke-style' formulation. Under various restrictions on its syntax, the Martini and Masini system captures K, $\mathsf{K4}$, T, and $\mathsf{S4}$, at the price of having one's terms annotated with indices. The Kripke-style formulation simplifies some of this presentation, but does not dispense with indices. A simpler Fitch-style system for K was extensively studied by Clouston [Clo18].

There is also a calculus in the style of Bierman and de Paiva for K, which was introduced by Bellin, de Paiva, and Ritter [BdPR01]. This suffered from some technical issues that were later mitigated by Kakutani [Kak07]. Some of the unsatisfactory aspects of this calculus are discussed by de Paiva and Ritter [dPR11]: they trace its roots to the aforementioned system of Bellin [Bel85] for GL, who hints at (but does not study) systems for K and $\mathsf{K4}$.

We argue that the dual-context formulations introduced in the present work lead to simpler calculi: our terms feature neither delayed substitutions, nor are they littered with indices. This simplifies the metatheory, and makes them more practicable. Unlike Bierman-de Paiva style calculi, dual-context calculi are simple enough to lead to implementations: see e.g. the experimental work of Wickline, Lee and Pfenning [WLP98] on metaprogramming with dual-context $\mathsf{S4}$. Moreover, our calculi are simple enough to enable large-scale pen-and-paper proofs: Shulman [Shu18] used a dual-context dependent type theory to produce a formal proof of Brouwer's fixed point theorem. Thus, the fact that dual-context style can extended to a range of sub-$\mathsf{S4}$ modalities is of independent interest in exploring applications in various sub-$\mathsf{S4}$ settings.

Nevertheless, we ought to stress one central detail: our systems strongly preserve products. That is, the proof theory of dual-context systems induces an isomorphism of types $\Box(A \times B) \cong \Box A \times \Box B$. In contrast, systems in the Bierman-de Paiva style prove a bi-implication that does *not* necessarily extend to an isomorphism. Indeed, the categorical semantics of Bierman-de Paiva $\mathsf{S4}$ require only a *lax monoidal comonad*, which comes with a natural transformation $\Box A \times \Box B \to \Box(A \times B)$ that is not necessarily invertible. This means our work is closer to the system of Clouston [Clo18], who requires that the modality have a left adjoint, and thus that it preserve products. All things considered, if we want to avoid product preservation we must revert to the system with delayed substitutions.

1. The Logics in Question

We are concerned with $(\wedge \to \Box)$ fragment of five of the most commonly encountered *normal modal logics* [BdRV01, §1.6] [Fit93, §§1.6–1.7] [HC96, §§2–3] [Boo94, §§4–5]: K (abbrv. CK), $\mathsf{K4}$ (abbrv. $\mathsf{CK4}$), T (abbrv. CT), GL (abbrv. CGL), and $\mathsf{S4}$ (abbrv. $\mathsf{CS4}$). In this section we shall discuss their common characteristics, and present a Hilbert system for each.

1.1. Constructive modal logics.

All of the above logics belong to the group of *constructive modal logics*. These are a family of intuitionistic modal logics which have been cherry-picked to satisfy a specific desideratum, namely to have a well-behaved, Gentzen-style proof theory, and thereby an associated computational interpretation.

The special behaviour of these logics is even more appreciable when the possibility modality $(\Diamond)$ is taken into consideration. First, the de Morgan duality between $\Box$ and $\Diamond$ breaks down, rendering them logically independent. For that reason we shall mostly refer to the $\Box$ as the *box* modality. Second, the principles $\Diamond(A \vee B) \to \Diamond A \vee \Diamond B$ and $\neg\Diamond\bot$ are not provable. These two principles are tautologies if we employ traditional Kripke semantics [Kri63]. Thus, the way to a computational interpretation seems to necessitates that we eschew the Kripkean analysis. Even though $\Diamond$ is essential in pinpointing the salient differences between constructive modal logics and other forms of intuitionistic modal logic—e.g. those studied by [Sim94]—it seems that its computational interpretation is not very crisp. Hence, we restrict our study to the better-behaved, and seemingly more applicable box modality.

1.2. Preliminaries.

All of our modal logics shall be inductively defined sets of formulæ—the *theorems* of the logic. These formulæ are generated by the Backus-Naur form

$$A, B \ ::= \ p_i \mid \top \mid A \wedge B \mid A \to B \mid \Box A$$

where p_i is drawn from a countable set of propositions. The sets of theorems will be generated by *axioms*, closed under some *inference rules*. The set of axioms will always contain (a) all the instances the axioms of intuitionistic propositional logic, but over modal formulæ; and (b) all instances of the *normality* axiom, also known as axiom K (after Kripke). The set of inference rules will contain rules for using axioms and assumptions, modus ponens, and the modal rule of *necessitation*, namely

$$\frac{A \in \mathcal{L}}{\Box A \in \mathcal{L}}$$

The only thing that will then vary between any two of our logics $\mathcal{L}$ will be the set of axioms.

1.3. Axioms.

We write $(A_1) \oplus \cdots \oplus (A_n)$ to mean the set of theorems that are derivable from all instances of the axioms $(A_1), \ldots, (A_n)$ under the aforementioned rules of axiom, assumption, modus ponens, and necessitation. Furthermore, we write $(\mathsf{IPL}_\Box)$ to mean the set of all instances of the axiom schemata of intuitionistic propositional logic, but over modal formulæ. We will use the following modal axiom schemata:

(K)	$\Box(A \to B) \to (\Box A \to \Box B)$	$(\mathsf{4})$	$\Box A \to \Box\Box A$
(T)	$\Box A \to A$	(GL)	$\Box(\Box A \to A) \to \Box A$

Constructive K is then defined to be the minimal *normal* constructive modal logic. Constructive $\mathsf{K4}$ adds axiom $\mathsf{4}$ to that. Likewise, constructive T is the result of adding axiom T

$$\frac{}{\Gamma, A, \Delta \vdash A}\,(\mathsf{assn}) \quad \frac{A \text{ is an axiom}}{\Gamma \vdash A}\,(\mathsf{ax}) \quad \frac{\Gamma \vdash A \to B \quad \Gamma \vdash A}{\Gamma \vdash B}\,(\mathsf{MP}) \quad \frac{\vdash A}{\Gamma \vdash \Box A}\,(\mathsf{nec})$$

Figure 1: Hilbert systems

to CK. Constructive $\mathsf{S4}$ results from mixing all these axiom schemas together. Finally, we obtain constructive GL from CK by adding the *Gödel-Löb* axiom GL. In summary:

$$\mathsf{CK} \stackrel{\text{def}}{=} (\mathsf{IPL}_\Box) \oplus (\mathsf{K}) \qquad \mathsf{CK4} \stackrel{\text{def}}{=} (\mathsf{IPL}_\Box) \oplus (\mathsf{K}) \oplus (\mathsf{4})$$

$$\mathsf{CT} \stackrel{\text{def}}{=} (\mathsf{IPL}_\Box) \oplus (\mathsf{K}) \oplus (\mathsf{T}) \qquad \mathsf{CGL} \stackrel{\text{def}}{=} (\mathsf{IPL}_\Box) \oplus (\mathsf{K}) \oplus (\mathsf{GL})$$

$$\mathsf{CS4} \stackrel{\text{def}}{=} (\mathsf{IPL}_\Box) \oplus (\mathsf{K}) \oplus (\mathsf{4}) \oplus (\mathsf{T})$$

1.4. **Hilbert systems.** We introduce a judgment of the form

$$\Gamma \vdash A$$

where Γ is a *context*, i.e. a list of formulæ defined by the grammar $\Gamma ::= \ \cdot \mid \Gamma, A$ where A is a single formula. We shall use the comma to also denote concatenation. For example, Γ, A, Δ shall stand for the juxtaposition of three things: the context Γ, the context consisting of the single formula A, and the context Δ.

This judgment is generated inductively, and includes rules for axioms and assumptions: see Figure 1. The main rule concerning the modality is that of *necessitation*, which we need state carefully. Otherwise, we risk invalidating the deduction theorem, leading to a common point of confusion in early work on the proof theory of modal logic: see [HN12] for a historical and technical account. To approach this issue, we need to recall that necessitation bears a likeness to universal quantification: $\Box A$ is a theorem just if A is a theorem, and there is no reason that this should be so if we need any assumptions to prove A. Thus, we should be able to infer $\Box A$ (under any assumptions) only if we can infer A without any assumptions at all. In symbols:

$$\frac{\vdash A}{\Gamma \vdash \Box A}$$

To indicate that we are using the Hilbert system for e.g. CK, we annotate the turnstile an write $\Gamma \vdash_{\mathsf{CK}} A$. We write $\Gamma \vdash A$ when the relevant statement pertains to all systems.

1.5. **Metatheory of Hilbert systems.**

1.5.1. *Structural rules.* We establish the following basic facts about all our Hilbert systems by a straightforward induction on the derivation of each premise.

Theorem 1.1 (Structural & Cut)**.** *The following rules are admissible.*[1]

(1) *(Weakening)* $\dfrac{\Gamma \vdash C}{\Gamma, A \vdash C}$ (2) *(Exchange)* $\dfrac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C}$

[1] Recall that a rule is *admissible* just if the existence of a proof of the antecedent implies the existence of a proof of the conclusion, where that existence is determined in our metatheory. In contrast, a rule is *derivable* just if a proof of the antecedent can be used verbatim as a constituent part of a proof of the conclusion.

(3) *(Contraction)* $\dfrac{\Gamma, A, A, \Delta \vdash C}{\Gamma, A, \Delta \vdash C}$ (4) *(Cut)* $\dfrac{\Gamma \vdash A \qquad \Gamma, A, \Delta \vdash C}{\Gamma, \Delta \vdash C}$

Theorem 1.2 (Deduction Theorem)**.** *The rule* $\dfrac{\Gamma, A \vdash B}{\Gamma \vdash A \to B}$ *is admissible.*

1.5.2. *Admissible Modal Rules.* We now consider some admissible rules that refer to the box modality. These will prove useful when we tackle the proof of equivalence between Hilbert systems and dual-context systems.

The first one is *Scott's rule*, which ensures that if we 'box' all our assumptions then we can 'box' the conclusion. We will see that, in categorical terms, Scott's rule expresses the fact that the box is a *functor*, and in particular one that *preserves products*. We write $\Box\Gamma$ to mean the context Γ which each assumption occurring in it boxed, i.e. $\Box(A_1, \ldots, A_n) \stackrel{\text{def}}{=} \Box A_1, \ldots, \Box A_n$.

Theorem 1.3 (Admissibility of Scott's rule)**.** *The following rule is admissible:*

$$\frac{\Gamma \vdash A}{\Box\Gamma \vdash \Box A}$$

Proof. Straightforward induction on the derivation of $\Gamma \vdash A$. We show the case for modus ponens. If the last step in the derivation of $\Gamma \vdash A$ is of the form

$$\frac{\begin{array}{c}\vdots\\ \Gamma \vdash B \to A\end{array} \qquad \begin{array}{c}\vdots\\ \Gamma \vdash B\end{array}}{\Gamma \vdash A}$$

then, by applying the induction hypotheses to the two subderivations, we obtain proofs of $\Box\Gamma \vdash \Box(B \to A)$ and $\Box\Gamma \vdash \Box B$. We can then use axiom K and modus ponens twice to build the desired proof:

$$\dfrac{\dfrac{\dfrac{}{\Box\Gamma \vdash \Box(B \to A) \to \Box B \to \Box A} \qquad \begin{array}{c}\vdots\\ \Box\Gamma \vdash \Box(B \to A)\end{array}}{\Box\Gamma \vdash \Box B \to \Box A} \qquad \begin{array}{c}\vdots\\ \Box\Gamma \vdash \Box B\end{array}}{\Box\Gamma \vdash \Box A}$$

□

The above theorem also follows by the deduction theorem. However, this proof implicitly considers derivations of $\Gamma \vdash A$ as terms of an underlying *modal combinatory logic*. It is an old observation by Curry and Feys [CF58] that combinatory logics vaguely correspond to Hilbert systems, and Pfenning [Pfe10] has sketched such a system for $\mathsf{CS4}$.

Next, we deal with a rule that is only derivable if the system contains the axiom T. The gist of the rule is that $\Box A$ is stronger than A, as it implies it in any context.

Theorem 1.4 (Admissibility of Veridicality)**.** *If $\mathcal{L} \in \{\mathsf{CT}, \mathsf{CS4}\}$, then the following rule is admissible:*

$$\frac{\Gamma \vdash_{\mathcal{L}} A}{\Box\Gamma \vdash_{\mathcal{L}} A}$$

Proof. By induction on the derivation of $\Gamma \vdash A$. All the cases are straightforward, except the assumption rule. If $\Gamma \vdash A$ because A occurs in Γ, then $\Box\Gamma \vdash \Box A$, and using modus ponens along with an instance of axiom T yields the result. $\Box$

Finally, we present a rule that we call the *Four rule*. As its name suggests, the Four rule encapsulates the deductive behaviour of axiom $\mathsf{4}$. In a nutshell, it expresses the fact that if something is derivable from $\Box\Box A$ then it is derivable from $\Box A$ itself.

The Four rule only pertains to logics that include all instances of $\mathsf{4}$. One of these logics is CGL, but in that case $\mathsf{4}$ is a theorem, so we begin by deriving it.

Lemma 1.5. $\vdash_{\mathsf{CGL}} \Box A \to \Box\Box A$

Proof. We follow [Boo94]. By using one of the conjunction axioms of ($\mathsf{IPL}_\Box$) and Scott's rule, we have $\Box(\Box A \wedge A) \vdash \Box A$, and hence $A, \Box(\Box A \wedge A) \vdash \Box A \wedge A$ by weakening, axiom, and the one of the conjunction axioms. Then, Scott's rule followed by the deduction theorem yield that $\Box A \vdash \Box\left(\Box\left(\Box A \wedge A\right) \to \Box A \wedge A\right)$. The conclusion matches the premise of the Gödel-Löb axiom, so using modus ponens gives yields $\Box A \vdash \Box\left(\Box A \wedge A\right)$. Cutting this with $\Box\left(\Box A \wedge A\right) \vdash \Box\Box A$ and using the deduction theorem completes the proof. $\Box$

Theorem 1.6 (Admissibility of the Four Rule)**.** *If $\mathcal{L}$ is a logic that includes $\mathsf{4}$ either as axiom or as theorem, i.e. if $\mathcal{L} \in \{\mathsf{CK4}, \mathsf{CGL}, \mathsf{CS4}\}$, then the following rule is admissible:*

$$\frac{\Box\Gamma, \Gamma \vdash_{\mathcal{L}} A}{\Box\Gamma \vdash_{\mathcal{L}} \Box A}$$

Proof. Induction on the derivation of $\Box\Gamma, \Gamma \vdash A$. Most cases are straightforward. If $\Box\Gamma, \Gamma \vdash A$ by the assumption rule, it follows that A either occurs in $\Box\Gamma$, or it occurs in Γ. If it occurs in $\Box\Gamma$, then it is of the form $\Box A'$; thus $\Box\Gamma \vdash \Box A'$, and using modus ponens alongside an instance of axiom $\mathsf{4}$ yields $\Box\Gamma \vdash \Box\Box A' = \Box A$. If, on the other hand, A occurs in Γ, then $\Box\Gamma \vdash \Box A$ by the assumption rule. $\Box$

A slightly weaker variant of the Four rule appears in [BdP00], and follows by weakening.

Corollary 1.7. *If $\mathcal{L} \in \{\mathsf{CK4}, \mathsf{CGL}, \mathsf{CS4}\}$, then the following rule is admissible:*

$$\frac{\Box\Gamma \vdash_{\mathcal{L}} A}{\Box\Gamma \vdash_{\mathcal{L}} \Box A}$$

If veridicality is admissible as well—i.e. in the case of $\mathsf{CS4}$—we can derive the theorem from the corollary. If $\Box\Gamma, \Gamma \vdash A$, then $\Box\Box\Gamma, \Box\Gamma \vdash A$ by veridicality, and repeatedly cutting with instances of $\Box B \vdash \Box\Box B$ yields $\Box\Gamma, \Box\Gamma \vdash A$. Repeated uses of exchange and contraction then show $\Box\Gamma \vdash A$, to which we apply the corollary.

Finally, we show that *Löb's rule* is admissible in CGL.

Theorem 1.8 (Löb's Rule)**.** *The rule* $\dfrac{\Box\Gamma, \Gamma, \Box A \vdash A}{\Box\Gamma \vdash \Box A}$ *is admissible in CGL.*

Proof. By the deduction theorem we infer that $\Box\Gamma, \Gamma \vdash \Box A \to A$, and hence, by the Four rule, $\Box\Gamma \vdash \Box(\Box A \to A)$. We then use the Gödel-Löb axiom and modus ponens. $\Box$

The following, which follows by weakening, is often quoted as Löb's rule.

Corollary 1.9. *The rule* $\dfrac{\Box\Gamma, \Box A \vdash A}{\Box\Gamma \vdash \Box A}$ *is admissible in CGL.*

2. From sequent calculi to dual contexts

We will now discuss the problems that one usually faces when devising modal λ-calculi for box modalities. We then demonstrate how the dual-context pattern decisively deals with many of these, by importing patterns found in well-behaved sequent calculi.

2.1. The perennial issues.

Most work on the subject is concentrated on essentially two kinds of calculi: those with *delayed substitutions*, following a style that was popularised by Bierman and de Paiva [BdP00]; and those employing *dual contexts*, a pattern that was imported into modal type theory by Davies and Pfenning [DP01, PD01].

Explicit substitutions à la Bierman & de Paiva. The calculus introduced by Bierman and de Paiva made use of a trick that was previously employed in the context of Intuitionistic Linear Logic by [BBdPH93] to ensure that substitution is admissible. The trick is simple: *if cut is not admissible, then we build it into the introduction rule.*

In the case of $\mathsf{CS4}$, the resultant syntax is an extension of the ordinary simply-typed λ-calculus. The extension is obtained by adding the following introduction rule:

$$\frac{\Gamma \vdash M_1 : \Box A_1 \quad \dots \quad \Gamma \vdash M_n : \Box A_n \qquad x_1 : \Box A_1, \dots, x_n : \Box A_n \vdash N : B}{\Gamma \vdash \mathsf{box}\ N\ \mathsf{with}\ M_1, \dots, M_n\ \mathsf{for}\ x_1, \dots x_n : \Box B}$$

In this example, $x_1, \dots x_n$ comprise all the free variables that may occur in N. They must all be 'modal,' in that their type has to start with a box. We are allowed to place a box in front of B, but we must provide a substitute M_i for each of these free variables. Of course, this M_i must also be of modal type. In short: *all the data that goes into the making of something of type $\Box B$ must be 'boxed.'* The given substitutes M_i are 'frozen' as part of the term of type $\Box B$: they become a *delayed substitution*[2] in the syntax. This is a combined introduction and cut rule: the introduction part ensures that modal data depend only on modal data, and the cut part ensures that substitution is admissible.

The elimination rule is simpler by comparison, and incorporates axiom T:

$$\frac{\Gamma \vdash M : \Box A}{\Gamma \vdash \mathsf{unbox}\ M : A}$$

In order to ensure admissibility of cut and hence subject reduction, the β-rule associated with these rules has the effect of unrolling the delayed substitutions *en masse*:

$$\mathsf{unbox}\ (\mathsf{box}\ N\ \mathsf{with}\ M_1, \dots, M_n\ \mathsf{for}\ x_1, \dots x_n) \longrightarrow N[M_1/x_1, \dots, M_n/x_n]$$

Calculi of this sort are notorious for suffering from two kinds of problems: the need for *commuting conversions*, and the lack of *'good' symmetries.*

- **Commuting Conversions:** In order to maintain the validity of vital proof-theoretic results, calculi with delayed subsitutions often require a large number of *commuting conversions*. The rôle of these conversions is to expose 'hidden' redexes, the existence of which spoil the so-called *subformula property*, i.e. the property that in normal proofs all detours have been eliminated. The issue of commuting conversions usually arises from positive connectives, such as disjunction and existence: see the book by Girard [GLT89, §10.1] for a particularly perspicuous discussion.

[2] These are often referred to as *explicit substitutions*. The present author reserves this term for those that are intentionally build into a calculus, and are not an artifact of proof-theoretic desires.

In calculi such as the above, commuting conversions invariably take the form of *structural rules* that reshuffle the delayed substitutions. Such rules are traditionally found in sequent calculi, but not in natural deduction, where they are often admissible. Their presence in a natural deduction system is incompatible with the view that natural deduction proofs comprise the "real proof objects"—see [GLT89, §5.4]. In the context of Bierman and de Paiva's system for $\mathsf{CS4}$, Goubault-Larrecq [GL96] argues that systems with such rules obscure the computational meaning of modal proofs.

'Good' symmetries: The calculus of Bierman and de Paiva for $\mathsf{CS4}$ exhibits reasonable symmetries: if we forget about the delayed substitutions for a moment, then we can see an introduction and an elimination rule, the latter post-inverse to the former: there is reasonable *harmony*.

Things are not that simple when it comes to other calculi of this sort. As a first example, consider the calculus of Bellin-de Paiva-Ritter [BdPR01] for CK. Its introduction rule is only slightly different to the one for $\mathsf{CS4}$, in that the free variables need not be of modal type. However, the substitutes for these free variables must still be modal. To wit:

$$\frac{\Gamma \vdash M_1 : \Box A_1 \quad \dots \quad \Gamma \vdash M_n : \Box A_n \quad x_1 : A_1, \dots, x_n : A_n \vdash N : B}{\Gamma \vdash \mathsf{box}\ N\ \mathsf{with}\ M_1, \dots, M_n\ \mathsf{for}\ x_1, \dots x_n : \Box B}$$

In this calculus there can be no harmony, for there is no elimination rule at all. Indeed, the only plausible 'β-rule' is very similar to a commuting conversion for $\mathsf{CS4}$ that was studied by Goubault-Larrecq [GL96]. Its function is to unbox any 'canonical' terms in the delayed substitutions; e.g

$$\mathsf{box}\ yx\ \mathsf{with}\ y, (\mathsf{box}\ M\ \mathsf{with}\ z\ \mathsf{for}\ z), N\ \mathsf{for}\ y, x, w \ \longrightarrow\ \mathsf{box}\ yM\ \mathsf{with}\ y, z, N\ \mathsf{for}\ y, z, w$$

This reduction locates 'boxed' delayed substitutions of 'boxed' proofs, and combines them into a single 'boxed' proof. See [Kak07] for a calculus for CK with this rule.

Secondly, this pattern leads to significant complexity in more complicated systems, e.g. when we need *diagonal assumptions*, which are natural in the case of GL. The only natural deduction system for CGL, which is due to Bellin [Bel85], is of this form. Translating the proof tree formulation to terms terms, its single modal rule is

$$\frac{\begin{array}{c}\Gamma \vdash M_1 : \Box B_1 \quad \dots \quad \Gamma \vdash M_m : \Box B_m \quad \Gamma \vdash N_1 : \Box C_1 \quad \dots \quad \Gamma \vdash N_n : \Box C_n \\ x_1 : B_1, \dots, x_m : B_m, y_1 : \Box C_1, \dots, y_n : \Box C_n, z : \Box A \vdash N : A\end{array}}{\Gamma \vdash \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ N\ \mathsf{with}\ M_1, \dots, M_n \mid N_1, \dots, N_n\ \mathsf{for}\ x_1, \dots x_m \mid y_1, \dots y_n : \Box A}$$

This calculus is virtually at the midpoint between the Bierman-de Paiva calculus for $\mathsf{CS4}$, and the Bellin-de Paiva-Ritter calculus for CK: some the assumptions that are being closed come are 'boxed,' and some are not. Evidently, this pattern is closely related to Löb's rule (Theorem 1.8) for CGL. It also features a *diagonal assumption* $z : \Box A$, which is bound in the resulting term.

Normalization for the terms of Bellin's calculus is by no means easy to describe. The only two small-step reductions that this calculus admits are also akin to commuting conversions. Using vector notation for succinctness, write $\mathsf{box}\ N\ \mathsf{with}\ \vec{M} \mid \vec{N}\ \mathsf{for}\ \vec{x} \mid \vec{y}$ to mean $\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ N\ \mathsf{with}\ \vec{M} \mid \vec{N}\ \mathsf{for}\ \vec{x} \mid \vec{y}$, where z is a fresh variable that is not free in N. Following Bellin, we call this a $\mathsf{K4R}$ application. The first small-step reduction ("$\mathsf{K4R}$

reduction") is essentially the one for CK given above, which here takes the form

$$\begin{aligned} & \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M\ \mathsf{with}\ \vec{P}, (\mathsf{box}\ N\ \mathsf{with}\ \vec{S} \mid \vec{T}\ \mathsf{for}\ \vec{v} \mid \vec{v'}), \vec{Q} \mid \vec{R}\ \mathsf{for}\ \vec{x}, w, \vec{z} \mid \vec{y} \\ \longrightarrow\ & \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M[N/w]\ \mathsf{with}\ \vec{P}, \vec{S}, \vec{Q} \mid \vec{T}, \vec{R}\ \mathsf{for}\ \vec{x}, \vec{v}, \vec{y} \mid \vec{v'}, \vec{y} \end{aligned}$$

The second rule ("segment reduction step") is closer to a commuting conversion for CS4:

$$\begin{aligned} & \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M\ \mathsf{with}\ \vec{P} \mid \vec{Q}, (\mathsf{fix}\ b\ \mathsf{in}\ N\ \mathsf{with}\ \vec{S} \mid \vec{T}\ \mathsf{for}\ \vec{v} \mid \vec{w}), \vec{R}\ \mathsf{for}\ \vec{x} \mid \vec{y}, z, \vec{y'} \\ \longrightarrow\ & \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M[(\mathsf{fix}\ b\ \mathsf{in}\ N\ \mathsf{with}\ \vec{c} \mid \vec{d}\ \mathsf{for}\ \vec{v} \mid \vec{w})/z]\ \mathsf{with}\ \vec{P}, \vec{S} \mid \vec{Q}, \vec{T}, \vec{R}\ \mathsf{for}\ \vec{x}, \vec{c} \mid \vec{y}, \vec{d}, \vec{y'} \end{aligned}$$

Normalization of proofs for Bellin's calculus is contingent on an auxiliary recursive algorithm, whose purpose is to turn every modal rule into a K4R application by *eliminating diagonal assumptions*, so that the small-step reductions can then simplify the remaining cuts. The details are far too complex to reproduce here.

It is thus evident that, once we step out of CS4, the use of systems based on 'mixed' introduction rules with delayed subsitutions becomes less and less tenable: the resulting systems lack harmony, and proof normalization becomes frighteningly complicated.

In order to reach a better solution we must overcome two problems: (a) we must 'decouple' the two flavours—introduction and cut—that together constitute the introduction or mixed modal rules; and (b) we must minimize as much as possible the commuting conversions—in particular, we should strive to free them from any computational content.

Dual contexts. The right intuition for achieving this decoupling was introduced by Girard [Gir93] in his attempt to combine classical, intuitionistic, and linear logic in one system, and also independently by Andreoli [And92] in the context of linear logic programming. The idea is simple and can be turned into a slogan: *segregate assumptions.* This means that we should divide our usual context of assumptions in two, or—even better—think of it as consisting of *two zones.* We should think of one zone as the *primary zone*, and the assumptions occuring in it as the 'ordinary' sort of assumptions. The other zone is the *secondary zone*, and the assumptions in it normally have a different flavour. In this setting the introduction rule explains the interaction between the two contexts, whereas the elimination rule effects substitution for the secondary context.

This idea has been most profitable in the case of the *Dual Intuitionistic Linear Logic* (DILL) of Plotkin and Barber [Plo93, Bar96] where the primary context consists of *linear* assumptions, and the secondary one consists of *intuitionstic* assumptions. The 'of course' modality (!) of Linear Logic is very much like a S4 modality, and—simply by lifting the linearity restrictions—Davies and Pfenning [PD01, DP01] adapted this work to the modal logic CS4 with considerable success. In this system, hereafter referred to as *dual constructive S4* (DS4), the primary context consists of *intuitionistic* assumptions, whereas the secondary context consists of *modal* assumptions.

The systems of Barber, Plotkin, Davies and Pfenning do not immediately seem adaptable to other logics. Indeed, the pattern may at first seem limited to modalities like 'of course' and the necessity of S4, which categorically correspond to *comonads.* As a comonad can be decomposed into an adjunction, one might think that the dual-context pattern implicitly makes use of the underlying universal property. In the rest of this section we show that not only this is not so, but that the dual-context style can be adapted to capture the necessity fragments of all of the logics introduced in §1.

2.2. **Deriving dual-context calculi.** Gentzen introduced the sequent calculus in the 1930s [Gen35a, Gen35b] in order to study normalization of proofs, which we call *cut elimination* in this context. A proof in the sequent calculus consists of a tree of *sequents*, which take the form $\Gamma \vdash A$, where Γ is a context. Thus in our notation a sequent is a different name for a judgment, like the ones in natural deduction.[3] The rules, however, are different: they come in two flavours: *left rules* and *right rules*. Broadly speaking, right rules are exactly the introduction rules of natural deduction, as they only concern the conclusion A of the sequent. The left rules play a rôle similar to that of elimination rules, but they do so by 'gerrymandering' with the assumptions in Γ. See the in-depth discussion of Girard [GLT89, §5.4] on the correspondence between natural deduction and sequent calculus.

The first attempts to forge sequent calculi for modal logics began in the 1950s, with the formulation of a sequent calculus for $\mathsf{S4}$ by Curry [Cur52] and Ohnisi and Matsumoto [OM57, OM59]. There was also some limited success for other simple modal logics, mainly involving the axioms we discuss here: see the surveys of Ono [Ono98] and Wansing [Wan02].

2.2.1. *The Introduction Rules.* Let us consider the (intuitionistic) right rule for the logic $\mathsf{S4}$:

$$\frac{\Box\Gamma \vdash A}{\Box\Gamma \vdash \Box A}\,(\Box\mathcal{R})$$

One cannot help but notice this rule has an intuitive computational interpretation in terms of 'flow of data.' We can read it as follows: if only modal data are used in inferring A, then we may obtain $\Box A$. Like in the Bierman-de Paiva calculus, only 'boxed' things may flow into something that is 'boxed.'

Let us at the same time consider dual-context judgments. These take the form

$$\Delta \,;\, \Gamma \vdash A$$

where both Δ and Γ are contexts. The assumptions in Δ are to be thought of as *modal*, whereas the assumptions in Γ are run-of-the-mill intuitionistic assumptions. A loose translation of such a a judgment to the 'ordinary' sort would be

$$\Delta \,;\, \Gamma \vdash A \qquad \rightsquigarrow \qquad \Box\Delta, \Gamma \vdash A$$

Under this translation, if we 'mimic' the right rule for $\mathsf{S4}$ we would obtain the following:

$$\frac{\Delta \,;\, \cdot \vdash A}{\Delta \,;\, \cdot \vdash \Box A}$$

where $\cdot$ denotes the empty context. However, natural deduction systems do not have any structural rules, so we have to include some kind of opportunity to weaken the context in the above rule. If we do so, the result is

$$\frac{\Delta \,;\, \cdot \vdash A}{\Delta \,;\, \Gamma \vdash \Box A}$$

Under the translation described above, this is exactly the right rule for $\mathsf{S4}$, with some extra weakening included. Incidentally, it is also exactly the introduction rule of Davies-Pfenning dual-context system for $\mathsf{S4}$ [PD01].

[3] Fundamental differences arise in the classical case, which features sequents of the form $\Gamma \vdash \Delta$ where Γ and Δ are lists of formulae. In intuitionistic logic Δ consists of at most one formula: see [GLT89, §5.1.3].

This pattern can be harvested to turn the right rules for the box in sequent calculi into introduction rules in dual-context systems. We tackle each case separately, except T, which we discuss in §2.2.3.

K. The case for K is slightly harder to fathom at first sight. This is because its sequent only has a single rule for the modality, namely *Scott's rule*:

$$\frac{\Gamma \vdash A}{\Box\Gamma \vdash \Box A}$$

As Bellin et al. [BdPR01] discuss, this rule is unsavoury: it is both a left and a right rule at the same time. It cannot be split into two rules, which is the pattern that bestows sequent calculus its fundamental symmetries. Despite this, Scott's rule is reasonably well-behaved. Leivant [Lei81] and Valentini [Val82] showed that incorporating it yields a system which admits cut elimination. It also also appears in the sequent calculus for CK studied by Wijesekera [Wij90].

According to the preceding translation, our introduction rule should be

$$\frac{\cdot\,;\Delta \vdash A}{\Delta\,;\cdot \vdash \Box A}$$

Indeed, we emulate Scott's rule by ensuring that *all the intuitionistic assumptions must become modal, at once.* The final form is reached again by adding opportunities for weakening:

$$\frac{\cdot\,;\Delta \vdash A}{\Delta\,;\Gamma \vdash \Box A}$$

At this point the reader may vehemently protest that this introduction rule is not in the spirit of natural deduction, as we are shamelessly messing with assumptions. So much is true. But it is also true that even the most well-behaved fragments of natural deduction are not really trees, but involve some 'back edges,' e.g. to record when and which assumptions are discharged: see [GLT89, §2.1]. The situation is even more involved when it comes to the not-so-well-behaved positive fragment ($\vee\,\exists$): for example, the elimination rule for $\vee$, namely

$$\frac{\Gamma \vdash A \vee B \quad \Gamma, A \vdash C \quad \Gamma, B \vdash C}{\Gamma \vdash C}$$

involves the silent elimination of two 'temporary assumptions,' A and B. Rules involving such temporary assumptions are known as rules in the style of Schroeder-Heister [SH84]. The sum of it all is this: *the proofs were never really trees.*

Consequently, the shameless shuffling of assumptions shall not be a cause for concern. In fact, there is a simple way to think about the 'jump' that the context Δ makes from intuitionistic to modal position. If we are writing down a deduction on the blackboard, and we wish to introduce a box in front of the conclusion, then all we have to do is to place a mark on all the assumptions that are open at that point. This does not discharge them, but it makes them modal: there shall be a fundamentally different way of substituting for them, and it shall be a little more complicated than the simple splicing of a proof tree at a leaf.

K4. The correct sequent calculus rule for the logic $\mathsf{K4}$, as well as the proof of cut elimination, is due to Sambin and Valentini [SV82]. Using elements from his joint work with Sambin, as well some counterexamples found in the work of Leivant [Lei81] on GL, Valentini noticed that the key property induced by axiom $\mathsf{4}$ is that anything derivable from $\Box\Box A$ is derivable from $\Box A$. The following (mixed left-and-right) rule for the encapsulates this insight:

$$\frac{\Box\Gamma, \Gamma \vdash A}{\Box\Gamma \vdash \Box A}$$

Thus, to derive $\Box A$ from a bunch of boxed assumptions, it suffices to derive A from two copies of the same assumptions, one boxed and one unboxed.[4] This co-occurence of the same assumptions in two forms will cause some mild technical complications in the next section, but that will clarify the structure of the 'flow of data' in $\mathsf{K4}$.

Following our previous recipe, a direct translation of this rule yields

$$\frac{\Delta \,;\, \Delta \vdash A}{\Delta \,;\, \Gamma \vdash \Box A}$$

GL. The correct formulation of sequent calculus for GL is a difficult problem that has repeatedly received attention. There are simple solutions that guarantee that we can derive all and only theorems of GL, but they fail to satisfy cut elimination.

The first attempt at a cut-free sequent calculus was that of Leivant [Lei81]. Soon thereafter, Valentini [Val83] showed that Leivant's proof of cut elimination was incorrect. Sambin and Valentini [SV80] describe a procedure for building cut-free proofs for all provable sequents, but their proof is semantic and goes through Kripke structures, and hence does not constitute Gentzen-style cut elimination. In [SV82], the same authors collect and describe in detail many early approaches, the reasons they do or do not work, and all relevant results. Finally, Valentini [Val83] shows that the same rule admits cut elimination, but the proof is rather complicated, and derives from the techniques of Bellin [Bel85]. Recent progress on clarifying that result may be found in Goré and Ramanayke [GR12]. Another approach, this time based on infinitary derivations, has been followed by Shamkanov [Sha14].

The Leivant-Valentini sequent calculus rule for GL is

$$\frac{\Box\Gamma, \Gamma, \Box A \vdash A}{\Box\Gamma \vdash \Box A}$$

The only difference between this rule and the rule for $\mathsf{K4}$ is the 'diagonal assumption' $\Box A$. We can straightforwardly use our translation to state it as an introduction rule:

$$\frac{\Delta \,;\, \Delta, \Box A \vdash A}{\Delta \,;\, \Gamma \vdash \Box A}$$

[4] Indeed, this is the Four rule we presented in §1.

2.2.2. *The Elimination Rule.* As discussed in §2.1, in a dual-context calculus we consider one context to be the *primary zone*, and the other to be the *secondary zone*. Assumptions in the primary zone are discharged by λ-abstraction. Thus, the function space of DILL is linear, whereas the function space of $\mathsf{DS4}$ is intuitionistic.

In contrast, substituting for assumptions in the secondary zone is the capacity of the *elimination rule*. This is a customary pattern for dual-context calculi: unlike primary assumptions, substitution for secondary assumptions is essentially a cut rule. In the term assignment system we will consider later, this takes the form of a *delayed substitution*, a type of 'let construct.' The rationale is this: the rest of the system controls how secondary assumptions arise and are used, and the elimination rule uniformly allows one to substitute for them.[5] To wit:

$$\frac{\Delta\,;\Gamma \vdash \Box A \quad \Delta, A\,;\Gamma \vdash C}{\Delta\,;\Gamma \vdash C}\,(\Box\mathcal{E})$$

The reader might protest that we are trying to pass a cut rule as an elimination rule. Notwithstanding the hypocrisy, this is not only common, but also the best presently known solution that recovers the patterns of introduction and elimination in the presence of modality. It is the core of our second slogan: *in dual-context systems, substitution is a cut rule for secondary assumptions.*

One cannot help but notice that such rules are also in the infamous style of Schroeder-Heister [SH84], and also similar to the elimination rule for disjunction. As we discussed in §2.1, this kind of rule is known to be problematic, as it automatically necessitates some commuting conversions: unavoidably, the conclusion C has no structural relationship with anything else in sight. The pressing question is whether this is an acceptable state of affairs. Unless we are to engage in more complicated and radical schemes, the present author is afraid that we must settle for it. Put simply, there is no good way to do away with commuting conversions: they are part-and-parcel of any sufficiently complicated type theory. All we can hope for is to minimize their number, and state them systematically.

2.2.3. *A second variable rule.* We have conveniently avoided discussing two things up to this point: the left rule for $\Box$ in $\mathsf{S4}$, which is the only one of our logics that has both left and right rules, and the case of T. These two are intimately related.

The left rule for necessity in $\mathsf{S4}$ is

$$\frac{\Gamma, A \vdash B}{\Gamma, \Box A \vdash B}\,(\Box\mathcal{L})$$

We can intuitively read it as follows: if A suffices to infer B, then $\Box A$ is more than enough. This is a form of the veridicality rule from §1.5.2, and encapsulates the axiom $\Box A \to A$. Together with Scott's rule, it forms a sequent calculus where cut is admissible; this is mentioned by Wansing [Wan02], and is attributed to Ohnisi and Matsumoto [OM57].

One way of emulating this rule in our framework would be to have a construct that makes an assumption 'jump' from one context to another, but that is inelegant and leads to an unworkable metatheory. The right way to imitate the left rule is to include a rule that allows one to use a *modal* assumption as if it were merely intuitionistic. To wit:

$$\frac{}{\Delta, A, \Delta'\,;\Gamma \vdash A}$$

[5] Alternative approaches have also been considered. For example, one could introduce another abstraction operator, i.e. a 'modal λ.' This has been adopted by Pfenning [Pfe01] in a dependently-typed setting.

$$
\begin{array}{rl}
\textbf{Types} & A, B \ ::= p_i \mid A \times B \mid A \to B \mid \Box A \\
\textbf{Typing Contexts} & \Gamma, \Delta \ ::= \cdot \mid \Gamma, x : A \\
\textbf{Terms} & M, N ::= x \mid \lambda x : A.\ M \mid MN \mid \langle M, N \rangle \mid \pi_i(M) \mid \\
& \qquad \mathsf{box}\ M \mid \mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N
\end{array}
$$

$$
\frac{}{\Delta\,;\Gamma, x : A, \Gamma' \vdash x : A}\ (\mathsf{var}) \qquad \frac{}{\Delta, u : A, \Delta'\,;\Gamma \vdash u : A}\ (\Box\mathsf{var})
$$

$$
\frac{\Delta\,;\Gamma \vdash M : A \quad \Delta\,;\Gamma \vdash N : B}{\Delta\,;\Gamma \vdash \langle M, N \rangle : A \times B}\ (\times\mathcal{I}) \qquad \frac{\Delta\,;\Gamma \vdash M : A_1 \times A_2}{\Delta\,;\Gamma \vdash \pi_i\,(M) : A_i}\ (\times\mathcal{E}_i)
$$

$$
\frac{\Delta\,;\Gamma, x : A \vdash M : B}{\Delta\,;\Gamma \vdash \lambda x : A.\ M : A \to B}\ (\to\mathcal{I}) \qquad \frac{\Delta\,;\Gamma \vdash M : A \to B \quad \Delta\,;\Gamma \vdash N : A}{\Delta\,;\Gamma \vdash MN : B}\ (\to\mathcal{E})
$$

$$
\frac{\cdot\,;\Delta \vdash M : A}{\Delta\,;\Gamma \vdash \mathsf{box}\ M : \Box A}\ (\Box\mathcal{I}_{\mathsf{K}}) \qquad \frac{\Delta\,;\Delta^{\perp} \vdash M^{\perp} : A}{\Delta\,;\Gamma \vdash \mathsf{box}\ M : \Box A}\ (\Box\mathcal{I}_{\mathsf{K4}})
$$

$$
\frac{\Delta\,;\Delta^{\perp}, z^{\perp} : \Box A \vdash M^{\perp} : A}{\Delta\,;\Gamma \vdash \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M : \Box A}\ (\Box\mathcal{I}_{\mathsf{GL}}) \qquad \frac{\Delta\,;\cdot \vdash M : A}{\Delta\,;\Gamma \vdash \mathsf{box}\ M : \Box A}\ (\Box\mathcal{I}_{\mathsf{S4}})
$$

$$
\frac{\Delta\,;\Gamma \vdash M : \Box A \qquad \Delta, u : A\,;\Gamma \vdash N : C}{\Delta\,;\Gamma \vdash \mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N : C}\ (\Box\mathcal{E})
$$

Rules for $\mathsf{S4}$: $(\Box\mathcal{I}_{\mathsf{S4}})$ and $(\Box\mathsf{var})$

Rules for T: $(\Box\mathcal{I}_{\mathsf{K}})$ and $(\Box\mathsf{var})$

Figure 2: Definition and typing judgments

This translates back to the sequent $\Box\Delta, \Box A, \Box\Delta', \Gamma \vdash A$, which follows by $(\Box\mathcal{L})$.

A rule like this was introduced by Plotkin and Barber [Plo93, Bar96] for *dereliction* in DILL, and is also essential in Davies and Pfenning's DS4. In our case, we use it in combination with the introduction rule for K in order to make a system for T.

3. Types, terms, and metatheory

We now collect all the observations we have made in order to turn our natural deduction systems into term assignment systems, i.e. typed λ-calculi. First, we annotate each assumption A with a variable, which we write as $x : A$. Then, we annotate each judgment $\Delta\,;\Gamma \vdash A$ with a term M representing the entire deduction that with that judgment as its conclusion—see [GLT89, §3] or [Gal93, SU06] for an introduction to term assignment.

The grammars defining types, terms and contexts, as well as the typing rules for all our systems can be found in Figure 2. All of our systems contain the introduction and elimination rules for products and functions, the variable rule (var), and the box elimination rule $(\Box\mathcal{E})$. Each of the systems for K, K4 and GL also contain the corresponding introduction rule, e.g. $(\Box\mathcal{I}_{\mathsf{K}})$. Finally, the systems for T and S4 each contain two additional rules: the modal variable rule $(\Box\mathsf{var})$, and a modal introduction rule—$(\Box\mathcal{I}_{\mathsf{K}})$ in the first case, and

$(\Box\mathcal{I}_{\mathsf{S4}})$ in the second. When we are at risk of confusion we annotate the turnstile with a subscript to indicate which system we mean.

From this point onwards, we assume Barendregt's conventions: terms are equal up to α-conversion, and bound variables are silently renamed whenever necessary. In $\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N$, u is a bound variable in N. Finally, we write $N[M/x]$ to mean capture-avoiding substitution of M for x in N. Furthermore, we shall assume that whenever we write a judgment like $\Delta\,;\Gamma \vdash M : A$, then Δ and Γ are *disjoint*, in the sense that $\textsc{Vars}\,(\Delta) \cap \textsc{Vars}\,(\Gamma) = \emptyset$, where $\textsc{Vars}\,(x_2 : A_1, \ldots, x_n : A_n) \stackrel{\text{def}}{=} \{x_1, \ldots, x_n\}$. This causes a mild technical complication in the cases $\mathsf{K4}$ and GL. Fortunately, the solution is relatively simple, and we explain it now.

3.1. **Complementary variables.** Naïvely annotating the rule for $\mathsf{K4}$ would yield

$$\frac{\Delta\,;\Delta \vdash M : A}{\Delta\,;\Gamma \vdash \mathsf{box}\ M : \Box A}$$

This, however, violates our convention that the two contexts are disjoint: the same variables will appear both at modal and intuitionistic positions. To overcome this we introduce the notion of *complementary variables*. Let $\mathcal{V}$ be the set of term variables for our calculi. A *complementation function* is an *involution* on variables. That is, it is a bijection $(-)^\perp : \mathcal{V} \xrightarrow{\cong} \mathcal{V}$ which happens to be its own inverse, i.e. $\left(x^\perp\right)^\perp = x$. The idea is that, if u is the modal variable representing some assumption in Δ, we will write $u^\perp$ to refer to a variable x, uniquely associated to u, and representing the same assumption, but without a box in front. For technical reasons, we would like that $x^\perp$ be the same variable as u.

We extend the involution to contexts:

$$(x_1 : A_1, \ldots, x_n : A_n)^\perp \stackrel{\text{def}}{=} x_1^\perp : A_1, \ldots, x_n^\perp : A_n$$

We also inductively extend $(-)^\perp$ to terms, with the exception that it must not change anything inside a $\mathsf{box}\ (-)$ construct. It also need not change any bound modal variables, as for $\mathsf{K4}$ and GL these shall only occur under $\mathsf{box}\ (-)$ constructs:

$$(\lambda x : A.M)^\perp \stackrel{\text{def}}{=} \lambda x^\perp : A.\ M^\perp \qquad (MN)^\perp \stackrel{\text{def}}{=} M^\perp N^\perp$$

$$\langle M, N\rangle \stackrel{\text{def}}{=} \langle M^\perp, N^\perp\rangle \qquad (\pi_i(M))^\perp \stackrel{\text{def}}{=} \pi_i(M^\perp)$$

$$(\mathsf{box}\ M)^\perp \stackrel{\text{def}}{=} \mathsf{box}\ M \qquad (\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N)^\perp \stackrel{\text{def}}{=} \mathsf{let\ box}\ u \Leftarrow M^\perp\ \mathsf{in}\ N^\perp$$

We use this machinery to modify the rule, so as to maintain disjoint contexts. When we encounter an introduction rule for the box and the context Δ gets 'copied' to the intuitionistic position, we will complement all variables in the copy, as well as all variables occuring in M, but not under any $\mathsf{box}\ (-)$ constructs, or bound by a let:

$$\frac{\Delta\,;\Delta^\perp \vdash M^\perp : A}{\Delta\,;\Gamma \vdash \mathsf{box}\ M : \Box A}$$

As an example, here is a derivation of $\cdot\,;\Box A \vdash \Box(A \wedge \Box A)$:

$$
\dfrac{\cdot\,; x : \Box A \vdash x : \Box A \qquad \dfrac{\dfrac{\dfrac{}{u : A\,; u^{\perp} : A \vdash u^{\perp} : A} \qquad \dfrac{\dfrac{}{u : A\,; u^{\perp} : A \vdash u^{\perp} : A}}{u : A\,; u^{\perp} : A \vdash \mathsf{box}\ u : \Box A}}{u : A\,; u^{\perp} : A \vdash \langle u^{\perp}, \mathsf{box}\ u\rangle : A \times \Box A}}{u : A\,; x : \Box A \vdash \mathsf{box}\ \langle u, \mathsf{box}\ u\rangle : \Box(A \times \Box A)}}{\cdot\,; x : \Box A \vdash \mathsf{let\ box}\ u \Leftarrow x\ \mathsf{in\ box}\ \langle u, \mathsf{box}\ u\rangle : \Box(A \times \Box A)}
$$

We extend complementation to finite sets of variables, by setting $\{x_1, \dots, x_n\} \stackrel{\text{def}}{=} x_1^{\perp}, \dots, x_n^{\perp}$. It is not hard to see that the involutive behaviour of $(-)^{\perp}$ is invariant inherited by these extensions, and that a number of operations commute with $(-)^{\perp}$.

Lemma 3.1.
(1) *For any context* Δ, $\left(\Delta^{\perp}\right)^{\perp} \equiv \Delta$.
(2) *For any finite set of variables* S, $\left(S^{\perp}\right)^{\perp} = S$.
(3) *For any context* Δ, $\textsc{Vars}\left(\Delta^{\perp}\right) = (\textsc{Vars}\,(\Delta))^{\perp}$.
(4) *If* S, T *are finite sets of variables, then* $S \subseteq T$ *implies* $S^{\perp} \subseteq T^{\perp}$.

There is a simple relationship between complementation and substitution:

Theorem 3.2. *If* $u^{\perp}$ *is not free in* M, *then* $(M[N/u])^{\perp} \equiv M^{\perp}[N, N^{\perp}/u, u^{\perp}]$.

Proof. By induction on M. Recall that $M \not\equiv u^{\perp}$ by assumption. The cases of λ-abstraction, application, pairing, projection, and $\mathsf{let\ box}\ u \Leftarrow (-)\ \mathsf{in}\ (-)$ follow by the IH.

Case(u). Then $(M[N/u])^{\perp} \equiv N^{\perp} \equiv u^{\perp}[N, N^{\perp}/u, u^{\perp}] \equiv M^{\perp}[N, N^{\perp}/u, u^{\perp}]$.

Case$(v \not\equiv u)$. Then $(M[N/u])^{\perp} \equiv v^{\perp} \equiv v^{\perp}[N, N^{\perp}/u, u^{\perp}] \equiv M^{\perp}[N, N^{\perp}/u, u^{\perp}]$.

Case$(\mathsf{box}\ M')$. We have:

$$\left(\mathsf{box}\ (M'[N/u])\right)^{\perp} \equiv \mathsf{box}\ (M'[N/u]) \equiv (\mathsf{box}\ M')^{\perp}[N, N^{\perp}/u, u^{\perp}]$$

where the last step follows because $\mathsf{box}\ M' \equiv (\mathsf{box}\ M')^{\perp}$, and $u^{\perp}$ is not free in M'. ∎

To conclude our discussion of complementary variables, we carefully define what it means for a pair of contexts to be well-defined.

Definition 3.3 (Well-defined contexts)**.** A pair of contexts $\Delta\,;\Gamma$ is *well-defined* just if
(1) They are *disjoint*, i.e. $\textsc{Vars}\,(\Delta) \cap \textsc{Vars}\,(\Gamma) = \emptyset$.
(2) In the cases of $\mathsf{K4}$ and GL, no two complementary variables occur in the same context; that is, $\textsc{Vars}\,(\Gamma) \cap \textsc{Vars}\left(\Gamma^{\perp}\right) = \emptyset$ and $\textsc{Vars}\,(\Delta) \cap \textsc{Vars}\left(\Delta^{\perp}\right) = \emptyset$.

The second condition is easy to enforce, and will prove useful in some technical proofs.

3.2. **Free variables: boxed and unboxed.**

Definition 3.4 (Free variables)**.**

(1) The *free variables* $\text{FV}(M)$ of a term M are defined by

$$\text{FV}(x) \stackrel{\text{def}}{=} \{x\} \qquad \text{FV}(MN) \stackrel{\text{def}}{=} \text{FV}(M) \cup \text{FV}(N)$$

$$\text{FV}(\lambda x : A.\ M) \stackrel{\text{def}}{=} \text{FV}(M) - \{x\} \qquad \text{FV}(\langle M, N\rangle) \stackrel{\text{def}}{=} \text{FV}(M) \cup \text{FV}(N)$$

$$\text{FV}(\pi_i(M)) \stackrel{\text{def}}{=} \text{FV}(M) \qquad \text{FV}(\mathsf{box}\ M) \stackrel{\text{def}}{=} \text{FV}(M)$$

$$\text{FV}(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N) \stackrel{\text{def}}{=} \text{FV}(M) \cup (\text{FV}(N) - \{u\})$$

and for GL we replace the clause for $\mathsf{box}\,(-)$ with $\text{FV}(\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M) \stackrel{\text{def}}{=} \text{FV}(M) - \{z\}$.

(2) The *unboxed free variables* $\text{FV}_0(M)$ of a term are those that *do not* occur under the scope of a $\mathsf{box}\,(-)$ construct. The formal definition involves replacing the clause of $\text{FV}(-)$ for $\mathsf{box}\,(-)$ with $\text{FV}_0(\mathsf{box}\ M) \stackrel{\text{def}}{=} \emptyset$, and, for GL, with $\text{FV}_0(\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M) \stackrel{\text{def}}{=} \emptyset$.

(3) The *boxed free variables* $\text{FV}_{\geq 1}(M)$ of a term M are those that *do* occur under the scope of a $\mathsf{box}\,(-)$ construct. The formal definition involves replacing the clauses of $\text{FV}(-)$ for variables and for $\mathsf{box}\,(-)$ by $\text{FV}_{\geq 1}(x) \stackrel{\text{def}}{=} \emptyset$ and $\text{FV}_{\geq 1}(\mathsf{box}\ M) \stackrel{\text{def}}{=} \text{FV}(M)$, and, for GL, by $\text{FV}_{\geq 1}(\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M) \stackrel{\text{def}}{=} \text{FV}(M) - \{z\}$.

We can then prove the following theorem by a simple induction on terms.

Theorem 3.5 (Free variables)**.**

(1) *If* $\Delta\,;\Gamma, x : A, \Gamma' \vdash M : A$ *and* $x \notin \text{FV}(M)$*, then* $\Delta\,;\Gamma, \Gamma' \vdash M : A$.
(2) *If* $\Delta, u : A, \Delta'\,;\Gamma \vdash M : A$ *and* $u \notin \text{FV}(M)$*, then* $\Delta, \Delta'\,;\Gamma \vdash M : A$.
(3) *For every term* M*,* $\text{FV}(M) = \text{FV}_0(M) \cup \text{FV}_{\geq 1}(M)$.
(4) *For every term* M*,* $\text{FV}_0\left(M^{\perp}\right) = \text{FV}_0(M)^{\perp}$.
(5) *For every term* M*,* $\text{FV}_{\geq 1}\left(M^{\perp}\right) = \text{FV}_{\geq 1}(M)$.
(6) *If* $\mathcal{S} \in \{\mathsf{DK}, \mathsf{DK4}, \mathsf{DGL}\}$ *and* $\Delta\,;\Gamma \vdash_{\mathcal{S}} M : A$*, then*

$$\text{FV}_0(M) \subseteq \text{VARS}(\Gamma), \qquad \text{FV}_{\geq 1}(M) \subseteq \text{VARS}(\Delta)$$

(7) *If* $\mathcal{S} \in \{\mathsf{DS4}, \mathsf{DT}\}$ *and* $\Delta\,;\Gamma \vdash_{\mathcal{S}} M : A$*, then*

$$\text{FV}_0(M) \subseteq \text{VARS}(\Gamma) \cup \text{VARS}(\Delta), \qquad \text{FV}_{\geq 1}(M) \subseteq \text{VARS}(\Delta)$$

3.3. **Structural theorems.** As expected, our systems satisfy the standard menu of structural results: weakening, contraction, exchange, and cut rules are admissible.

Theorem 3.6 (Structural & Cut)**.** *The following rules are admissible in all systems:*

(1) *(Weakening)*

$$\frac{\Delta\,;\Gamma, \Gamma' \vdash M : C}{\Delta\,;\Gamma, x : A, \Gamma' \vdash M : C}$$

(2) *(Exchange)*

$$\frac{\Delta\,;\Gamma, x : A, y : B, \Gamma' \vdash M : C}{\Delta\,;\Gamma, y : B, x : A, \Gamma' \vdash M : C}$$

(3) *(Contraction)*

$$\frac{\Delta\,;\Gamma, x : A, y : A, \Gamma' \vdash M : A}{\Delta\,;\Gamma, w : A, \Gamma' \vdash M[w, w/x, y] : A}$$

(4) *(Cut)*

$$\frac{\Delta\,;\Gamma \vdash N : A \qquad \Delta\,;\Gamma, x : A, \Gamma' \vdash M : C}{\Delta\,;\Gamma, \Gamma' \vdash M[N/x] : C}$$

Proof. By induction on the typing derivation of M. As an example, we show the case of $(\Box\mathcal{I}_{\mathsf{K}})$ for weakening. Suppose $\Delta\,;\Gamma, \Gamma' \vdash M : C$ by $(\Box\mathcal{I}_{\mathsf{K}})$. Then $M \equiv \mathsf{box}\ M'$ and $C \equiv \Box C'$ and $\cdot\,;\Delta \vdash M' : C'$. A single use of $(\Box\mathcal{I}_{\mathsf{K}})$ then yields $\Delta\,;\Gamma, x : A, \Gamma' \vdash M : C$. $\Box$

Theorem 3.7 (Modal Structural)**.** *The following rules are admissible:*

(1) *(Modal Weakening)*

$$\frac{\Delta, \Delta' \,;\, \Gamma \vdash M : C}{\Delta, u : A, \Delta' \,;\, \Gamma \vdash M : C}$$

(2) *(Modal Exchange)*

$$\frac{\Delta, x : A, y : B, \Delta' \,;\, \Gamma \vdash M : C}{\Delta, y : B, x : A, \Delta' \,;\, \Gamma \vdash M : C}$$

(3) *(Modal Contraction)*

$$\frac{\Delta, x : A, y : A, \Delta' \,;\, \Gamma \vdash M : C}{\Delta, w : A, \Delta' \,;\, \Gamma \vdash M[w, w/x, y] : C}$$

Proof. Straightforward induction on the derivation of the premise. As an example, we discuss a few cases of $(\Box\mathcal{I})$ for weakening, the rest being similar.

If $\Delta, \Delta' \,;\, \Gamma \vdash M : A$ by $(\Box\mathcal{I}_{\mathsf{K}})$, then $M \equiv \mathsf{box}\ N$ and $A \equiv \Box B$, with $\cdot\,;\, \Delta, \Delta' \vdash N : B$. We use Theorem 3.6 to obtain $\cdot\,;\, \Delta, x : A, \Delta' \vdash N : B$, and then apply $(\Box\mathcal{I}_{\mathsf{K}})$.

If $\Delta, \Delta' \,;\, \Gamma \vdash M : A$ by $(\Box\mathcal{I}_{\mathsf{K4}})$, then $M \equiv \mathsf{box}\ N$ and $A \equiv \Box B$, with $\Delta, \Delta' \,;\, \Delta^{\perp}, \Delta'^{\perp} \vdash N^{\perp} : B$. By the IH, we have that $\Delta, u : A, \Delta' \,;\, \Delta^{\perp}, \Delta'^{\perp} \vdash N^{\perp} : B$. We use Theorem 3.6 to deduce that $\Delta, u : A, \Delta' \,;\, \Delta^{\perp}, u^{\perp} : A, \Delta'^{\perp} \vdash N^{\perp} : B$, and then apply $(\Box\mathcal{I}_{\mathsf{K4}})$. □

Theorem 3.8 (Modal Cut)**.** *The following rules are admissible:*

(1) *(Modal Cut for* DK*)*

$$\frac{\cdot\,;\, \Delta \vdash_{\mathsf{DK}} N : A \quad \Delta, u : A, \Delta' \,;\, \Gamma \vdash_{\mathsf{DK}} M : C}{\Delta, \Delta' \,;\, \Gamma \vdash_{\mathsf{DK}} M[N/u] : C}$$

(2) *(Modal Cut for* $\mathsf{DK4}$*)*

$$\frac{\Delta \,;\, \Delta^{\perp} \vdash_{\mathsf{DK4}} N^{\perp} : A \quad \Delta, u : A, \Delta' \,;\, \Gamma \vdash_{\mathsf{DK4}} M : C}{\Delta, \Delta' \,;\, \Gamma \vdash_{\mathsf{DK4}} M[N/u] : C}$$

(3) *(Modal Cut for* DGL*)*

$$\frac{\Delta \,;\, \Delta^{\perp}, z^{\perp} : \Box A \vdash_{\mathsf{DGL}} N^{\perp} : A \quad \Delta, u : A, \Delta' \,;\, \Gamma \vdash_{\mathsf{DGL}} M : C}{\Delta, \Delta' \,;\, \Gamma \vdash_{\mathsf{DGL}} M\left[N\left[\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ N/z\right]/u\right] : C}$$

(4) *(Modal Cut for* $\mathsf{DS4}$*)*

$$\frac{\Delta \,;\, \cdot \vdash_{\mathsf{DS4}} N : A \quad \Delta, u : A, \Delta' \,;\, \Gamma \vdash_{\mathsf{DS4}} M : C}{\Delta, \Delta' \,;\, \Gamma \vdash_{\mathsf{DS4}} M[N/u] : C}$$

(5) *(Modal Cut for* DT*)*

$$\frac{\cdot\,;\, \Delta \vdash_{\mathsf{DT}} N : A \quad \Delta, u : A, \Delta' \,;\, \Gamma \vdash_{\mathsf{DT}} M : C}{\Delta, \Delta' \,;\, \Gamma \vdash_{\mathsf{DT}} M[N/u] : C}$$

Proof. By induction on the typing derivation of M. We show the case of $(\Box\mathcal{I})$, and—for $\mathsf{DS4}$ and DT—the case of modal variables $(\Box\mathsf{var})$.

(1) (DK) If $\Delta, u : A, \Delta' \,;\, \Gamma \vdash M : C$ by $(\Box\mathcal{I}_{\mathsf{K}})$, then $M \equiv \mathsf{box}\ M'$, $C \equiv \Box C'$, and

$$\cdot\,;\, \Delta, u : A, \Delta' \vdash M' : C'$$

By Theorem 3.6, we have

$$\cdot\,;\, \Delta, \Delta' \vdash M'[N/u] : C$$

and hence $\Delta, \Delta' ; \Gamma \vdash \mathsf{box}\ (M'[N/u]) : \Box C' \equiv C$ by an application of $(\Box\mathcal{I}_{\mathsf{K}})$. But

$$\mathsf{box}\ \left(M'[N/u]\right) \equiv \left(\mathsf{box}\ M'\right)[N/u] \equiv M[N/u]$$

and hence we have the result.

(2) $(\mathsf{DK4})$ If $\Delta, u : A, \Delta' ; \Gamma \vdash M : C$ by $(\Box\mathcal{I}_{\mathsf{K4}})$, then $M \equiv \mathsf{box}\ M'$, $C \equiv \Box C'$, and

$$\Delta, u : A, \Delta' ; \Delta^{\perp}, u^{\perp} : A, \Delta'^{\perp} \vdash M'^{\perp} : C'$$

By the IH, we have

$$\Delta, \Delta' ; \Delta^{\perp}, u^{\perp} : A, \Delta'^{\perp} \vdash M'^{\perp}[N/u] : C'$$

and by Theorem 3.6, that yields

$$\Delta, \Delta' ; \Delta^{\perp}, \Delta'^{\perp} \vdash M'^{\perp}[N, N^{\perp}/u, u^{\perp}] : C'$$

But, by Theorem 3.2, we have that $M'^{\perp}[N, N^{\perp}/u, u^{\perp}] \equiv (M'[N/u])^{\perp}$, and hence by a use of $(\Box\mathcal{I}_{\mathsf{K4}})$, we have

$$\Delta, \Delta' ; \Gamma \vdash \mathsf{box}\ \left(M'[N/u]\right) : \Box C' \equiv C$$

and hence the result.

(3) (DGL) If $\Delta, u : A, \Delta' ; \Gamma \vdash M : C$ by $(\Box\mathcal{I}_{\mathsf{GL}})$, then $M \equiv \mathsf{fix}\ y\ \mathsf{in}\ \mathsf{box}\ M'$, $C \equiv \Box C'$, and

$$\Delta, u : A, \Delta' ; \Delta^{\perp}, u^{\perp} : A, \Delta'^{\perp}, y^{\perp} : \Box C' \vdash M'^{\perp} : C'$$

Write $N_* \stackrel{\text{def}}{=} N[\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ N/z]$. By the first premise and the IH, we have that

$$\Delta, \Delta' ; \Delta^{\perp}, u^{\perp} : A, \Delta'^{\perp}, y^{\perp} : \Box C' \vdash M'^{\perp}\,[N_*/u] : C'$$

We now need to substitute for $u^{\perp}$. By an application of $(\Box\mathcal{I}_{\mathsf{GL}})$ to the first premise we have

$$\Delta ; \Delta^{\perp} \vdash \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ N : \Box A$$

and hence by Theorem 3.6 we substitute this into the first premise itself to get

$$\Delta ; \Delta^{\perp} \vdash N^{\perp}[\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ N/z^{\perp}] : A$$

But $N_*^{\perp} \equiv N^{\perp}[\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ N/z^{\perp}]$, so by weakening and Theorem 3.6, we obtain

$$\Delta, \Delta' ; \Delta^{\perp}, \Delta'^{\perp}, y^{\perp} : \Box C \vdash M'^{\perp}[N_*, N_*^{\perp}/u, u^{\perp}] : C'$$

But by well-definedness of contexts, $u^{\perp} \notin \mathrm{Fv}\,(M)$, so by Theorem 3.2 we have that $M'^{\perp}[N_*, N_*^{\perp}/u, u^{\perp}] \equiv (M'[N_*/u])^{\perp}$, and hence by a use of $(\Box\mathcal{I}_{\mathsf{GL}})$, we have

$$\Delta, \Delta' ; \Gamma \vdash \mathsf{fix}\ y\ \mathsf{in}\ \mathsf{box}\ \left(M'[N_*/u]\right) : \Box C' \equiv C$$

and hence the result.

(4) $(\mathsf{DS4})$

- If $\Delta, u : A, \Delta' ; \Gamma \vdash M : C$ by $(\Box\mathcal{I}_{\mathsf{S4}})$ then $M \equiv \mathsf{box}\ M'$ and $C \equiv \Box C'$ with

  $$\Delta, u : A, \Delta' ; \cdot \vdash M' : C$$

  The IH then yields $\Delta, \Delta' ; \cdot \vdash M'[N/u] : C$, and a single use of $(\Box\mathcal{I}_{\mathsf{S4}})$ yields the result.
- If $\Delta, u : A, \Delta' ; \Gamma \vdash M : C$ by $(\Box\mathsf{var})$ then $M \equiv v$ for some v such that $(v : C) \in \Delta, u : A, \Delta'$. There are two cases:
  - $u \equiv v$: then $M[N/u] \equiv N$ and $A \equiv C$. The premise $\Delta ; \cdot \vdash N : A$ along with weakening for both contexts yields the result.
  - $u \not\equiv v$: then $M[N/u] \equiv M$, and u does not occur in M. It is easy to show that if $\Delta, u : A, \Delta' ; \Gamma \vdash M : C$ and $u \notin \mathrm{Fv}_{\geq 1}\,(M)$ then $\Delta, \Delta' ; \Gamma \vdash M : C$.

(5) (DT)
- If $\Delta, u : A, \Delta' ; \Gamma \vdash M : C$ by $(\Box\mathcal{I}_{\mathsf{K}})$ then we proceed as in the case of DK.
- If $\Delta, u : A, \Delta' ; \Gamma \vdash M : C$ by $(\Box\mathsf{var})$ then $M \equiv v$ for some v such that $v : C \in \Delta, u : A, \Delta'$. There are two cases:
  - $u \equiv v$: then $M[N/u] \equiv N$ and $A \equiv C$. The premise $\cdot ; \Delta \vdash N : A$ along with Theorem 3.9 yields $\Delta ; \cdot \vdash N : A$. A series of weakenings for both contexts then yields the result.
  - $u \not\equiv v$: then $M[N/u] \equiv M$, and u does not occur in M. It is easy to show that if $\Delta, u : A, \Delta' ; \Gamma \vdash M : C$ and $u \notin \mathrm{FV}_{\geq 1}(M)$ then $\Delta, \Delta' ; \Gamma \vdash M : C$.

$\square$

Finally, in the cases where the T axiom is present, we may move variables from the intuitionstic to the modal context:

Theorem 3.9 (Modal Dereliction)**.** *If* $\mathcal{S} \in \{\mathit{DS4}, \mathit{DT}\}$, *then* $\dfrac{\Delta ; \Gamma, \Gamma' \vdash_{\mathcal{S}} M : A}{\Delta, \Gamma ; \Gamma' \vdash_{\mathcal{S}} M : A}$ *is admissible.*

Proof. By induction on the derivation of $\Delta ; \Gamma, \Gamma' \vdash M : A$. Most cases are straightfoward, except (var) and $(\Box\mathcal{I}_{\mathsf{S4}})/(\Box\mathcal{I}_{\mathsf{K}})$. If the judgment holds by (var), then $M \equiv x$ for some $(x : A) \in \Gamma, \Gamma'$. If $(x : A) \in \Gamma$, we use $(\Box\mathsf{var})$ to conclude that $\Delta, \Gamma ; \Gamma' \vdash x : A$. If $(x : A) \in \Gamma'$, then use of (var). If the judgment holds by $(\Box\mathcal{I}_{\mathsf{S4}})$ then $M \equiv \mathsf{box}\ M'$ and $A \equiv \Box A'$ for some M', A' with $\Delta ; \cdot \vdash M' : A'$. Repeated use of weakening for the modal context followed by an application of $(\Box\mathcal{I}_{\mathsf{S4}})$ yields the result. The case of $(\Box\mathcal{I}_{\mathsf{K}})$ is similar, but uses weakening for the intuitionistic context. $\square$

3.4. **Equivalence between Hilbert and dual systems.**

In this section we prove that our dual-context λ-calculi correspond to the Hilbert systems given in §1. This ties the knot with respect to the Curry-Howard correspondence.

Modulo the appearance of proof terms, the translation under which this equivalence is shown is the same one that we used in §2:

$$\Delta ; \Gamma \vdash_{\mathsf{D}\mathcal{L}} M : A \quad \rightsquigarrow \quad \Box\hat{\Delta}, \hat{\Gamma} \vdash_{\mathcal{L}} A$$

We write $\hat{\Gamma}$ to mean the context Γ with all the variables removed: if $\Gamma \equiv x_1 : A_1, \dots, x_n : A_n$, then $\hat{\Gamma} \stackrel{\text{def}}{=} A_1, \dots, A_n$.

One direction of the proof involves showing that the axioms are indeed derivable in the dual-context systems. The other direction involves showing the admissibility of the dual-context rules in the Hilbert systems.

First and foremost, we need to show that axiom (K) is derivable. It is easy to check that the term

$$\mathsf{ax}_{\mathsf{K}} \stackrel{\text{def}}{=} \lambda f : \Box(A \to B).\ \lambda x : \Box A.\ \mathsf{let\ box}\ g \Leftarrow f\ \mathsf{in\ let\ box}\ y \Leftarrow x\ \mathsf{in\ box}\ (g\, y)$$

has type $\Box(A \to B) \to \Box A \to \Box B$ in all our systems other than GL. For GL, we instead use

$$\mathsf{ax}_{\mathsf{K}}^{\mathsf{DGL}} \stackrel{\text{def}}{=} \lambda f : \Box(A \to B).\ \lambda x : \Box A.\ \mathsf{let\ box}\ g \Leftarrow f\ \mathsf{in\ let\ box}\ y \Leftarrow x\ \mathsf{in\ fix}\ z\ \mathsf{in\ box}\ (g\, y)$$

It is also not hard to see that in DK4 and DS4 the terms

$$\mathsf{ax}_4 \stackrel{\text{def}}{=} \lambda x : \Box A.\ \mathsf{let\ box}\ y \Leftarrow x\ \mathsf{in\ box}\ (\mathsf{box}\ y)$$

have type $\Box A \to \Box\Box A$, which is exactly axiom 4.

In the case of DGL, we need to show that the term

$$\mathsf{ax}_{\mathsf{GL}} \stackrel{\text{def}}{=} \lambda x : \Box(\Box A \to A).\ \mathsf{let\ box}\ f \Leftarrow x\ \mathsf{in}\ \ (\mathsf{fix}\ z\ \mathsf{in\ box}\ \ (f\ z))$$

has type $\Box(\Box A \to A) \to \Box A$. The most interesting part of the derivation can be found in Figure 3.

$$\dfrac{\dfrac{}{\cdots \vdash x : \Box(\Box A \to A)} \qquad \dfrac{\begin{array}{c}\vdots \\ \Delta, f : \Box A \to A\,;\Delta, f^{\perp} : \Box A \to A, z^{\perp} : \Box A \vdash f^{\perp}\ z^{\perp} : A\end{array}}{\Delta, f : \Box A \to A\,;\Gamma, x : \Box(\Box A \to A) \vdash \mathsf{fix}\ z\ \mathsf{in\ box}\ \ (f\ z) : \Box A}}{\dfrac{\Delta\,;\Gamma, x : \Box(\Box A \to A) \vdash \mathsf{let\ box}\ f \Leftarrow x\ \mathsf{in}\ \ (\mathsf{fix}\ z\ \mathsf{in\ box}\ \ (f\ z)) : \Box(\Box A \to A) \to \Box A}{\Delta\,;\Gamma \vdash \lambda x : \Box(\Box A \to A).\ \mathsf{let\ box}\ f \Leftarrow x\ \mathsf{in}\ \ (\mathsf{fix}\ z\ \mathsf{in\ box}\ \ (f\ z)) : \Box(\Box A \to A) \to \Box A}}$$

Figure 3: Derivation of the Gödel-Löb axiom in DGL

Finally, in DT and $\mathsf{DS4}$, the term

$$\mathsf{ax}_{\mathsf{T}} \stackrel{\text{def}}{=} \lambda x : \Box A.\ \mathsf{let\ box}\ y \Leftarrow x\ \mathsf{in}\ y$$

has type $\Box A \to A$, i.e. inhabits axiom T.

With all that we can show:

Theorem 3.10 (Hilbert to Dual)**.** *If Γ is a well-defined context and $\hat{\Gamma} \vdash_{\mathcal{L}} A$, then there exists a term M such that $\cdot\,;\Gamma \vdash_{\mathsf{D}\mathcal{L}} M : A$.*

Proof. By induction on the derivation of $\hat{\Gamma} \vdash_{\mathcal{L}} A$. In the case of the assumption rule, we use (var) to type the associated variable in $\hat{\Gamma}$. The cases for axioms of ($\mathsf{IPL}_{\Box}$) are easy. For the modal axioms, we use the terms derived above. For modus ponens we use application.

This leaves the case of necessitation. Suppose $\hat{\Gamma} \vdash_{\mathcal{L}} A$ by it; then $A \equiv \Box A'$, and $\vdash_{\mathcal{L}} A'$. By the IH, there is a term M' such that $\cdot\,;\cdot \vdash_{\mathsf{D}\mathcal{L}} M' : A'$. We then use the appropriate introduction rule for box—e.g. ($\Box\mathcal{I}_{\mathsf{K}}$), and so on—to obtain $\cdot\,;\Gamma \vdash_{\mathsf{D}\mathcal{L}} \mathsf{box}\ M' : \Box A'$. $\Box$

The essence of the opposite direction lies in showing that the rules of the dual-context calculus are admissible in the corresponding Hilbert system—that is, after erasing the proof terms. We have done most of the required work in §1.5.2.

Theorem 3.11 (Dual to Hilbert)**.** *If $\Delta\,;\Gamma \vdash_{\mathsf{D}\mathcal{L}} M : A$ then $\Box\hat{\Delta}, \hat{\Gamma} \vdash_{\mathcal{L}} A$.*

Proof. By induction on the derivation of $\Delta\,;\Gamma \vdash_{\mathsf{D}\mathcal{L}} M : A$.

If the premise holds by (var), then we use the assumption rule of the Hilbert system. If the last step in the derivation of the premise is the rule ($\to \mathcal{I}$), we use the IH followed by the Deduction Theorem (Theorem 1.2). If the last step is by ($\to \mathcal{E}$), we use modus ponens. It is simple to translate the rules that pertain to the product, namely ($\times\mathcal{I}$) and ($\times\mathcal{E}_i$) to uses of the IPL axioms pertaining to the product along with modus ponens. It is also not hard to see that, under the given translation, ($\Box\mathcal{E}$) can also be matched by a use of the IH along with an invocation of the admissibility of cut for Hilbert systems (Theorem 1.1). Uses of the modal variable rule ($\Box\mathsf{var}$) can be imitated by a use of the assumption rule, modus ponens, and an instance of the T axiom.

This leaves the introduction rules for the box. The rule ($\Box\mathcal{I}_{\mathsf{K}}$) is matched with Scott's rule (Theorem 1.3). The rule ($\Box\mathcal{I}_{\mathsf{K4}}$) is matched with the Four rule (Theorem 1.6). The rule

($\Box\mathcal{I}_{\mathsf{GL}}$) is matched with the generalized Löb rule (Theorem 1.8). Finally, the rule ($\Box\mathcal{I}_{\mathsf{S4}}$) is matched with the corollary to the Four rule (Corollary 1.7). $\square$

4. Reduction

We will now show that it is possible to eliminate cuts from proofs in our systems. This elimination of cuts will be complete, in the sense that normal proofs will satisfy the *subformula property*, i.e. they will not reference any external logical formulæ that are unrelated to their assumptions or conclusion. We will achieve this with the traditional technique of a confluent and strongly normalizing small-step reduction. Rather strikingly, this reduction will be the same across all systems—with the exception of DGL, whose term former for the introduction of the modality has a very different shape. This avoids repeating work to deal with different systems, as most of our proofs are by induction on the typing judgments, and most rules are shared between all systems. We discuss GL separately in §4.5.

In this paper we stop short of deciding equality of proofs, which is a much more challenging problem. There are many reasons that make it so. The first one is related to the known problematic behaviour of η-contraction. The second arises from the fact that our reduction does not eliminate all redundancy from our proofs. For example, the terms

$$\cdot\,;x:\Box A, y:B\vdash y:B \qquad\qquad \cdot\,;x:\Box A, y:B\vdash \mathsf{let\ box}\ u \Leftarrow x\ \mathsf{in}\ y:B$$

will be equal in the equational theory of §6.1, but will also be normal forms with respect to reduction. In a sense, there have to be additional commuting conversions that—amongst other things—'garbage collect' unnecessary eliminations, and which we only discover when we consider the categorical semantics §6. The third problem is deeper: it arises because $\Box$ essentially behaves as *positive connective*. Deciding equality in the presence of such connectives requires either advanced rewriting or categorical techniques. For example, the case of $\beta\eta$-equivalence in the presence of sums and an empty type was open until Scherer resolved it in 2017 [Sch17]; see *op. cit.* for an extensive bibliography.

Normalization of proofs for this kind of system has not been extensively studied before. Pfenning and Davies [PD01] hint at our notion, and use a strict subset of it as operational semantics [DP01]. A similar notion was studied in the context of DILL by Ohta and Hasegawa [OH06], including η-contractions and the full set of commuting conversions.

4.1. **The reduction.**

The notion of reduction $\longrightarrow$ is defined as the least relation satisfying the rules of Figure 4. This includes the usual β-reduction, plus the modal β-reduction

$$\mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ M\ \mathsf{in}\ N \longrightarrow N[M/u]$$

which is suggested by Theorem 3.8. It also includes congruences—so that reductions can happen anywhere in a term—and, finally, three *commuting conversions*, which are required for the subformula property to hold.

We begin with the following lemma, which shall also prove useful in §4.4. Recall the definition and discussion of complementary variables from §3.1.

Lemma 4.1 (Complement reduction)**.** *If $\Delta\,;\Delta^{\perp}\vdash_{\mathsf{DK4}} M^{\perp}:A$ then $M\longrightarrow N$ implies $M^{\perp}\longrightarrow N^{\perp}$.*

Proof. By induction on $M\longrightarrow N$. We only prove the cases that are not straightforward.

$$\frac{}{(\lambda x{:}A.\ M)N \longrightarrow M[N/x]} \quad \frac{}{\pi_i(\langle M_1, M_2\rangle) \longrightarrow M_i} \quad \frac{M \longrightarrow N}{\pi_i(M) \longrightarrow \pi_i(N)} \quad \frac{M_i \longrightarrow N_i \quad M_{1-i} \equiv N_{1-i}}{\langle M_0, M_1\rangle \longrightarrow \langle N_0, N_1\rangle}$$

$$\frac{M \longrightarrow N}{\lambda x{:}A.\ M \longrightarrow \lambda x{:}A.\ N} \quad \frac{M \longrightarrow N}{\mathsf{box}\ M \longrightarrow \mathsf{box}\ N} \quad \frac{M \longrightarrow N}{MP \longrightarrow NP} \quad \frac{P \longrightarrow Q}{MP \longrightarrow MQ}$$

$$\frac{M \longrightarrow N}{\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ P \longrightarrow \mathsf{let\ box}\ u \Leftarrow N\ \mathsf{in}\ P} \quad \frac{P \longrightarrow Q}{\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ P \longrightarrow \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ Q}$$

$$\frac{}{\pi_i\left(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N\right) \longrightarrow \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ \pi_i(N)}$$

$$\frac{}{\left(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ P\right) Q \longrightarrow \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ PQ}$$

$$\frac{}{\mathsf{let\ box}\ v \Leftarrow (\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N)\ \mathsf{in}\ P \longrightarrow \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ \mathsf{let\ box}\ v \Leftarrow N\ \mathsf{in}\ P}$$

Figure 4: Reduction

Case(β for $\to$). Suppose $\Delta\,;\Delta^{\perp}, x : B \vdash M : C$ for some B, C. Then

$$((\lambda x : B.\ M)N)^{\perp} \equiv (\lambda x^{\perp} : B.\ M^{\perp})N^{\perp} \longrightarrow M^{\perp}[N^{\perp}/x^{\perp}]$$

We want to show that the latter is just $(M[N/x])^{\perp}$. We can rename x so that it does not clash with any variable nor its complement in the context of M. We will show that (a) $x^{\perp} \notin \mathrm{Fv}\,(M)$, and that (b) $x \notin \mathrm{Fv}\left(M^{\perp}\right)$. By (a) we may apply Theorem 3.2 and then use (b) to infer that

$$(M[N/x])^{\perp} \equiv M^{\perp}[N, N^{\perp}/x, x^{\perp}] \equiv M^{\perp}[N^{\perp}/x^{\perp}]$$

Both (a) and (b) follow from Theorem 3.5. For (a): by (3) it suffices to show that $x^{\perp} \notin \mathrm{Fv}_0\,(M)$ and $x^{\perp} \notin \mathrm{Fv}_{\geq 1}\,(M)$. But $x^{\perp}$ does not occur in either context, so we use (6). For (b): we know that $x^{\perp} \notin \mathrm{Fv}_0\,(M)$, so this implies by (4) that $x \notin \mathrm{Fv}_0\left(M^{\perp}\right)$. Hence, it suffices by (3) to also show that $x \notin \mathrm{Fv}_{\geq 1}\left(M^{\perp}\right)$. But by (5) the latter is equal to $\mathrm{Fv}_{\geq 1}\,(M)$, and by well-formedness of contexts we know that know that $x \notin \mathrm{Vars}\,(\Delta)$, so by (6) it is not in $\mathrm{Fv}_{\geq 1}\,(M)$ either.

Case(β for $\mathsf{box}\ (-)$). It is easy to see that

$$\begin{aligned}
(\mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ M\ \mathsf{in}\ N)^{\perp} &\equiv \mathsf{let\ box}\ u \Leftarrow (\mathsf{box}\ M)^{\perp}\ \mathsf{in}\ N^{\perp} \\
&\equiv \mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ M\ \mathsf{in}\ N^{\perp} \\
&\longrightarrow N^{\perp}[M/u]
\end{aligned}$$

It now suffices to show that (a) $u^{\perp} \notin \mathrm{Fv}\,(N)$, and that (b) $u^{\perp} \notin \mathrm{Fv}\left(N^{\perp}\right)$. For then Theorem 3.2 applies, and $(N[M/u])^{\perp} \equiv N^{\perp}[M, M^{\perp}/u, u^{\perp}] \equiv N^{\perp}[M/u]$. We will use Theorem 3.5 again, recalling that $\Delta, u : A\,; \Delta^{\perp} \vdash N^{\perp} : A$.

For (a): by (3) it suffices to show that $u^{\perp} \notin \mathrm{Fv}_0\,(N)$ and $u^{\perp} \notin \mathrm{Fv}_{\geq 1}\,(N)$. By (4) and (5) it suffices to show that $u \notin \mathrm{Fv}_0\left(N^{\perp}\right)$ and $u^{\perp} \notin \mathrm{Fv}_{\geq 1}\left(N^{\perp}\right)$. Both follow by (6), as neither u nor $u^{\perp}$ are allowed to occur anywhere in Δ and $\Delta^{\perp}$.

For (b): we know that $u^{\perp} \notin \mathrm{Fv}_{\geq 1}\left(N^{\perp}\right)$, so by (3) it suffices to show that $u^{\perp} \notin \mathrm{Fv}_0\left(N^{\perp}\right)$. But we can use (6): u cannot be in Δ, so $u^{\perp}$ cannot be in $\Delta^{\perp}$.

□

We then show that

Theorem 4.2 (Subject reduction)**.** *If $\Delta\,;\Gamma \vdash M : A$ and $M \longrightarrow N$, then $\Delta\,;\Gamma \vdash N : A$.*

Proof. By induction on $M \longrightarrow N$. Most cases follow straightforwardly from the IH. The cases for the β rules follow from Theorems 3.6 and 3.8. □

4.2. **Subformula property.** A calculus satisfies the subformula property when any *normal* proof (i.e. one that has no reducts) of a formula A from assumptions Γ only involves formulæ that are either (a) subformulæ of the conclusion of A, or (b) subformulæ of some premise in Γ. This is tantamount to saying that the proof has a very specific structure: it proceeds by eliminating logical symbols of assumptions in Γ, and then uses the results to construct a proof of A using only introduction rules. In short, *the proof has no detours*, and proceeds as quickly as possible from assumptions to conclusion: see [Pra65] and [GLT89].

Without the commuting conversions of Figure 4, our systems do not satisfy the subformula property. The reason is the presence of the elimination rule

$$\frac{\Delta\,;\Gamma \vdash M : \Box A \qquad \Delta, u : A\,;\Gamma \vdash N : C}{\Delta\,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N : C}\ (\Box\mathcal{E})$$

Notice that the conclusion C is given to us by the *minor premise* $\Delta, u : A\,;\Gamma \vdash N : C$, and it is structurally unrelated to $\Box A$, the *major premise* that is being eliminated: in Girard's terminology, it is *parasitic*. This is so because—as we discussed in §2—the elimination rule is secretly a kind of *cut rule*, or a rule in the style of Schroeder-Heister [SH84].

It is not so easy to see where the actual trouble lies at first. The point is that the $\mathsf{let\ box}\ u \Leftarrow (-)\ \mathsf{in}\ (-)$ construct may 'hide redexes.' Once we introduce the extra reductions that are needed, and prove the subformula property, this will become quite clear. But—in the meantime—let us consider three examples.

Suppose that $\Delta, u : A\,;\Gamma \vdash \langle N_1, N_2\rangle : A_1 \times A_2$ and $\Delta\,;\Gamma \vdash M : \Box A$ are normal forms, and that the latter is not of the form $\mathsf{box}\ (-)$. We may use $(\Box\mathcal{E})$ to obtain

$$\Delta\,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ \langle N_1, N_2\rangle : A_1 \times A_2$$

This is indeed—and should be!—a normal form. But what if we just want to prove A_1? We may apply one of the elimination rules for products to get

$$\Delta\,;\Gamma \vdash \pi_1\,(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ \langle N_1, N_2\rangle) : A_1$$

This is now a proof of A_1, but it surreptitiously contains a proof N_2 of A_2, which is entirely unrelated to A_1 (neither need be a subexpression of the other). The problem is that the eliminator $\mathsf{let\ box}\ u \Leftarrow (-)\ \mathsf{in}\ (-)$ obstructs the meeting of the destructor $\pi_1(-)$ with the constructor $\langle N_1, N_2\rangle$. The solution is to allow a commuting conversion that allows the two to meet by pulling the let construct outside:

$$\pi_1\,(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ \langle N_1, N_2\rangle) \longrightarrow \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ \pi_1(\langle N_1, N_2\rangle)$$

A similar situation occurs when $\Delta, u : A\,;\Gamma \vdash \lambda x : B.\ P : B \to C$. We may form

$$\Delta\,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ \lambda x : B.\ P : B \to C$$

which is a perfectly reasonable normal form. But if $\Delta\,;\Gamma \vdash Q : B$ then

$$\Delta\,;\Gamma \vdash (\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ \lambda x : B.\ P)\,Q : C$$

is not: we should be able to reduce

$$(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ \lambda x : B.\ P)\, Q \longrightarrow \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ (\lambda x : B.\ P)Q$$

Finally, there is third, less visible case of this phenomenon. If we understand $(\Box\mathcal{E})$ to be a 'bad' elimination, we have considered the cases of 'good' elimination ($\pi_i(-)$, application) following 'bad' elimination. The final case is that of 'bad' elimination following another 'bad' elimination. To give an example, let us consider an elimination after a $\mathsf{box}\ (-)$ introduction:

$$\Delta\,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in\ box}\ N : \Box A$$

We can then plug this into a term $\Delta, v : A\,;\Gamma \vdash P : C$ by eliminating the box:

$$\Delta\,;\Gamma \vdash \mathsf{let\ box}\ v \Leftarrow (\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in\ box}\ N)\ \mathsf{in}\ P : C$$

Now things are clear: the second let-construct is obstructing the meeting of the first let-construct with the introduction form $\mathsf{box}\ N$. We need to convert

$$\mathsf{let\ box}\ v \Leftarrow (\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in\ box}\ N)\ \mathsf{in}\ P \longrightarrow \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in\ let\ box}\ v \Leftarrow \mathsf{box}\ N\ \mathsf{in}\ P$$

whilst taking care to not confuse our bound variables.

With these commuting conversions in place we can now prove the property. We only need to slightly strengthen the induction hypothesis, using the notion of *principal branch.*

Theorem 4.3 (Subformula Property)**.** *Let $\Delta\,;\Gamma \vdash M : A$, and suppose M is a normal form.*

(1) *Every type occuring in the derivation of $\Delta\,;\Gamma \vdash M : A$ is either a subexpression of the type A, or a subexpression of a type in Δ or Γ.*
(2) *If M is an elimination form that is* not *form $\mathsf{let\ box}\ u \Leftarrow P\ \mathsf{in}\ Q$—i.e. if it is a projection $\pi_i(N)$ or an application PQ—then it entirely consists of a sequence of eliminations: that is, there is a sequence of types $A_0, \ldots, A_n$ for which*
   - *A_0 occurs in either Δ or Γ,*
   - *A_n is A, and*
   - *A_i is the* major premise *of an elimination whose conclusion is A_{i+1} for $0 \leq i < n$.*

   This sequence is called a principal branch. *In particular, A_n is a subexpression of A_0.*

Proof. By induction on the derivation of $\Delta\,;\Gamma \vdash M : A$. We omit some cases, as they are similar to others: the case of ($\Box\mathsf{var}$) is like that of an ordinary variable, the other introduction forms ($\langle M, N\rangle$, $\mathsf{box}\ M$) are similar to the case for λ, and the case for $\pi_i(M)$ is similar to the one for MN.

Case(x). Then $\Delta\,;\Gamma \vdash x : A$ and hence $(x : A) \in \Gamma$. This is the complete derivation, and satisfies both desiderata.

Case($\lambda x : A.\ M$). Then the immediate premise is of the form $\Delta\,;\Gamma, x : A \vdash M : B$. By the IH, all types that occur in that are either subexpressions of types in Δ or Γ, subexpressions of A, or subexpressions of B. All of these are subexpressions of types in Δ, Γ, or $A \to B$.

Case(MN). Then the major premise is $\Delta\,;\Gamma \vdash M : B \to A$ and the minor premise is $\Delta\,;\Gamma \vdash N : B$ for some type B.

Look at M: it cannot be a λ-abstraction, for that would make MN a redex. It also cannot be any other introduction rule, for it introduce some other type (e.g. $A \times B$ or $\Box A$). Hence, it must be an elimination. Of the eliminations, it cannot be a let-expression, for a commuting conversion would make that a redex too.

It follows that M is a 'good' elimination: either $\pi_i(-)$ or PQ. We can thus apply (2) from the IH thesis to conclude that there is a principal branch beginning with an assumption in Δ or Γ, and ending with $B \to A$. We can extend that to a principal branch for M, ending with A. This proves (2), and furthermore implies that $B \to A$ is a subexpression of some premise in either Δ or Γ.

Over to (1): applying the IH to the major premise we know that every type that occurs in the derivation of $\Delta\,;\Gamma \vdash M : B \to A$ is either a subexpression of a type in Δ or Γ, or a subexpression of $B \to A$. But we have already deduced that $B \to A$ is a subexpression of some premise in either Δ or Γ, so that all types occuring in the derivation of the major premise satisfy the desideratum.

Applying the IH to the minor premise, every type that occurs in the derivation of $\Delta\,;\Gamma \vdash N : B$ is either a subexpression of some type in Δ or Γ, or a subexpression of B. But B is a subexpression of $B \to A$, which in turn is a subexpression of a premise in one of the contexts. Hence all types occurring in that branch also occur in either Δ or Γ. This concludes the proof of this case, for we have examined all types appearing in the derivation.

Case($\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N$). The major premise is then $\Delta\,;\Gamma \vdash M : \Box B$ and the minor premise is $\Delta, u : B\,;\Gamma \vdash N : A$ for some B. (2) does not apply to let-constructs, so we only need to show (1).

Consider M. It cannot be a $\mathsf{box}\ (-)$, for that would make the entire term a redex. It also cannot be any other introduction form, because it would introduce a type of a different shape. Hence, it must hence be an elimination form, but not another let-construct, for that would be a redex too due to our commuting conversion. Therefore it must be either of the form $\pi_i(M')$ or of the form PQ. It follows that (2) of the IH applies: there is a principal branch beginning with a premise and ending with $\Box B$. In particular, $\Box B$ is a subexpression of some type in Δ or Γ.

By the IH, any type that occurs in the derivation of the major premise is either a subexpression of a type in Δ or Γ, or a subexpression of $\Box B$. But $\Box B$ is a subexpression of some type in one of those two contexts, so every type that occurs in the derivation of the major premise is actually a subexpression of a type in Δ or Γ.

As for the minor premise, any type that occurs in it is either a subexpression of a type in Δ or Γ, or a subexpression of the types B or A. But B is a subexpression of $\Box B$, which by our previous reasoning is in turn a subexpression of some type in either Δ or Γ. Thus all types occuring in it are either subexpressions of some type in Δ or Γ, or subexpressions of A. This concludes the proof. $\square$

4.3. **Confluence.** We will prove that

Theorem 4.4. *The reduction relation $\longrightarrow$ is confluent.*

One can show this result in many ways. We will use the method of *parallel reduction*, which was discovered by Tait and Martin-Löf. The history of the method and a few variations of it are discussed by Takahashi [Tak95]. The idea is simple: we introduce a second notion of reduction, $\Longrightarrow$, which we will 'sandwich' between reduction proper and its transitive closure, so that $\longrightarrow\ \subseteq\ \Longrightarrow\ \subseteq\ \longrightarrow^*$. We will then show that $\Longrightarrow$ has the diamond property. By the above inclusions, the transitive closure $\Longrightarrow^*$ of $\Longrightarrow$ is then equal to $\longrightarrow^*$, and hence $\longrightarrow$ is confluent. In fact, we will follow [Tak95] in doing something better: we will define

$$
\frac{}{x \Longrightarrow x}
\qquad
\frac{M \Longrightarrow N \quad P \Longrightarrow Q}{(\lambda x : A.\ M)P \Longrightarrow N[Q/x]}
\qquad
\frac{M_i \Longrightarrow N}{\pi_i(\langle M_1, M_2 \rangle) \Longrightarrow N}
\qquad
\frac{M \Longrightarrow N}{\pi_i(M) \Longrightarrow \pi_i(N)}
$$

$$
\frac{M_1 \Longrightarrow N_1 \quad M_2 \Longrightarrow N_2}{\langle M_1, M_2 \rangle \Longrightarrow \langle N_1, N_2 \rangle}
\qquad
\frac{M \Longrightarrow N}{\lambda x : A.\ M \Longrightarrow \lambda x : A.\ N}
\qquad
\frac{M \Longrightarrow N}{\mathsf{box}\ M \Longrightarrow \mathsf{box}\ N}
\qquad
\frac{M \Longrightarrow N \quad P \Longrightarrow Q}{MP \Longrightarrow NQ}
$$

$$
\frac{M \Longrightarrow N \quad P \Longrightarrow Q}{\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ P \Longrightarrow \mathsf{let\ box}\ u \Leftarrow N\ \mathsf{in}\ Q}
\qquad
\frac{M \Longrightarrow N \quad P \Longrightarrow Q}{\mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ P\ \mathsf{in}\ M \Longrightarrow N[Q/u]}
$$

$$
\frac{M \Longrightarrow N \quad P \Longrightarrow Q}{\pi_i\,(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ P) \Longrightarrow \mathsf{let\ box}\ u \Leftarrow N\ \mathsf{in}\ \pi_i(Q)}
$$

$$
\frac{P \Longrightarrow Q \quad M \Longrightarrow N \quad S \Longrightarrow T}{(\mathsf{let\ box}\ u \Leftarrow P\ \mathsf{in}\ M)\,S \Longrightarrow \mathsf{let\ box}\ u \Leftarrow Q\ \mathsf{in}\ NT}
$$

$$
\frac{P \Longrightarrow Q \quad M \Longrightarrow N \quad S \Longrightarrow T}{(\mathsf{let\ box}\ u \Leftarrow (\mathsf{let\ box}\ v \Leftarrow P\ \mathsf{in}\ M)\ \mathsf{in}\ S) \Longrightarrow \mathsf{let\ box}\ v \Leftarrow Q\ \mathsf{in}\ \mathsf{let\ box}\ u \Leftarrow N\ \mathsf{in}\ T}
$$

Figure 5: Parallel Reduction

for each term M its *complete development* $M^\star$. The complete development is intuitively defined by 'unrolling' all the redexes of M at once. We will then show that if $M \Longrightarrow N$, then $N \Longrightarrow M^\star$. $M^\star$ will then suffice to close the diamond:

$$
\begin{tikzcd}
& M \arrow[dl, Rightarrow] \arrow[dr, Rightarrow] & \\
P \arrow[dr, Rightarrow, dotted] & & Q \arrow[dl, Rightarrow, dotted] \\
& M^\star &
\end{tikzcd}
$$

The parallel reduction $\Longrightarrow$ is defined in Figure 5. It is immediate that

Lemma 4.5. $\Longrightarrow$ *is reflexive.*

Definition 4.6 (Complete development)**.** The *complete development* $M^\star$ of a term M is defined by the following clauses:

$$
\begin{array}{cc}
x^\star \stackrel{\text{def}}{=} x & (\langle M, N \rangle)^\star \stackrel{\text{def}}{=} \langle M^\star, N^\star \rangle \\
(\pi_i(\langle M_1, M_2 \rangle))^\star \stackrel{\text{def}}{=} M_i^\star & (\pi_i(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N))^\star \stackrel{\text{def}}{=} \mathsf{let\ box}\ u \Leftarrow M^\star\ \mathsf{in}\ \pi_i(N^\star) \\
((\lambda x : A.\ M)\,N)^\star \stackrel{\text{def}}{=} M^\star[N^\star/x] & ((\mathsf{let\ box}\ u \Leftarrow P\ \mathsf{in}\ M)\,N)^\star \stackrel{\text{def}}{=} \mathsf{let\ box}\ u \Leftarrow P\ \mathsf{in}\ M^\star N^\star \\
(\pi_i(M))^\star \stackrel{\text{def}}{=} \pi_i(M^\star) & (\lambda x : A.\ M)^\star \stackrel{\text{def}}{=} \lambda x : A.\ M^\star \\
(\mathsf{box}\ M)^\star \stackrel{\text{def}}{=} \mathsf{box}\ M^\star & (\mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ M\ \mathsf{in}\ N)^\star \stackrel{\text{def}}{=} N^\star[M^\star/u] \\
(MN)^\star \stackrel{\text{def}}{=} M^\star N^\star & (\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N)^\star \stackrel{\text{def}}{=} \mathsf{let\ box}\ u \Leftarrow M^\star\ \mathsf{in}\ N^\star
\end{array}
$$

$$
(\mathsf{let\ box}\ u \Leftarrow (\mathsf{let\ box}\ v \Leftarrow P\ \mathsf{in}\ M)\ \mathsf{in}\ N)^\star \stackrel{\text{def}}{=} \mathsf{let\ box}\ v \Leftarrow P^\star\ \mathsf{in}\ \mathsf{let\ box}\ u \Leftarrow M^\star\ \mathsf{in}\ N^\star
$$

First, a little lemma capturing the essence of parallel reduction:

Lemma 4.7. *If* $M \Longrightarrow N$ *and* $P \Longrightarrow Q$, *then* $M[P/x] \Longrightarrow N[Q/x]$.

Proof. Straightforward induction on $M \Longrightarrow N$. ∎

And here is the main result:

Theorem 4.8. *If $M \Longrightarrow P$, then $P \Longrightarrow M^\star$.*

Proof. By induction on $M \Longrightarrow P$. The case for variables is trivial, the case for the congruence rules follows from the IH, and β for function types is as usual. We show the rest.

Case(β for $\times$). Then we have $\pi_i(\langle M_1, M_2\rangle) \Longrightarrow M_i'$, with $M_i \Longrightarrow M_i'$. By the IH, $M_i' \Longrightarrow M_i^\star \equiv (\pi_i(\langle M_1, M_2\rangle))^\star$.

Case(β for $\Box$). Then we have $\mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ M\ \mathsf{in}\ N \Longrightarrow N'[M'/u]$ where $M \Longrightarrow M'$ and $N \Longrightarrow N'$. By the IH, $M' \Longrightarrow M^\star$ and $N' \Longrightarrow N^\star$. It follows that $N'[M'/u] \Longrightarrow N^\star[M^\star/u] \equiv (\mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ M\ \mathsf{in}\ N)^\star$ by Lemma 4.7.

Case(comm. conv. for $\times$). Then we have $\pi_i(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ P) \Longrightarrow \mathsf{let\ box}\ u \Leftarrow N\ \mathsf{in}\ \pi_i(Q)$ where $M \Longrightarrow N$ and $P \Longrightarrow Q$. By the IH, $N \Longrightarrow M^\star$ and $Q \Longrightarrow P^\star$, whence $\mathsf{let\ box}\ u \Leftarrow N\ \mathsf{in}\ \pi_i(Q) \Longrightarrow \mathsf{let\ box}\ u \Leftarrow M^\star\ \mathsf{in}\ \pi_i(P^\star)$, which is α-equivalent to $(\pi_i(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ P))^\star$.

The cases for the other commuting conversions are similar. $\Box$

4.4. **Strong normalization.** In this section, we will prove that

Theorem 4.9. *The reduction relation $\longrightarrow$ is strongly normalizing.*

There is a very orderly way of doing so for all systems, save for GL. The idea is to *embed* the modal proofs in the simply-typed λ-calculus, for which strong normalization is a known result [Pra71, dG02]. This strategy is used by [MM96, TI10], and hinted at for the dual-context $\mathsf{S4}$ system by [DP01]. Because of the binding structure of $\mathsf{let\ box}\ u \Leftarrow (-)\ \mathsf{in}\ (-)$, one cannot do so by simply erasing the modalities: we would then map modal β-reductions to syntactic equality in the simply-typed λ-calculus, which would not provide enough leverage to lift strong normalization to the modal calculi.

Instead, we use a strategy inspired by the proof of strong normalization for Moggi's monadic metalanguage [BBdP98]: we will interpret $\Box$ as the *product-with-the-unit comonad.*[6] More specifically, we define a translation $(-)^{\maltese}$ of modal types to simple types:

$$(p_i)^{\maltese} \stackrel{\text{def}}{=} p_i \qquad (A \times B)^{\maltese} \stackrel{\text{def}}{=} A^{\maltese} \times B^{\maltese}$$

$$(A \to B)^{\maltese} \stackrel{\text{def}}{=} A^{\maltese} \to B^{\maltese} \qquad (\Box A)^{\maltese} \stackrel{\text{def}}{=} \mathbf{1} \times A^{\maltese}$$

Next, we extend $(-)^{\maltese}$ to terms:

$$(x)^{\maltese} \stackrel{\text{def}}{=} x \qquad (\langle M, N\rangle)^{\maltese} \stackrel{\text{def}}{=} \langle M^{\maltese}, N^{\maltese}\rangle$$

$$(\pi_i(M))^{\maltese} \stackrel{\text{def}}{=} \pi_i(M^{\maltese}) \qquad (\lambda x : A.\ M)^{\maltese} \stackrel{\text{def}}{=} \lambda x : A^{\maltese}.\ M^{\maltese}$$

$$(MN)^{\maltese} \stackrel{\text{def}}{=} M^{\maltese} N^{\maltese} \qquad (\mathsf{box}\ M)^{\maltese} \stackrel{\text{def}}{=} \langle *, M^{\maltese}\rangle$$

$$(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N)^{\maltese} \stackrel{\text{def}}{=} N^{\maltese}[\pi_2(M^{\maltese})/u]$$

where $* : \mathbf{1}$ is the introduction form for the unit type. We can then show that

Theorem 4.10 (Simulation)**.**

[6][BBdP98] translate the monadic type TA to the exception monad $\mathbf{1} + (-)$. They simulate the commuting conversions for 'bind' using the commuting conversions for coproducts, which leads to a more direct proof.

(1) $(M[N/x])^{\maltese} \equiv M^{\maltese}[N^{\maltese}/x]$
(2) *If* $\Delta \,;\Gamma \vdash_{\mathsf{D}\mathcal{L}} M : A$ *then* $\Delta^{\maltese}, \Gamma^{\maltese} \vdash M^{\maltese} : A^{\maltese}$ *in the simply-typed* λ*-calculus.*
(3) *If* $M \longrightarrow N$ *then* $M^{\maltese} \longrightarrow^*_\beta N^{\maltese}$ *in the simply-typed* λ*-calculus.*

Proof.
(1) By induction on M.
(2) By induction on $\Delta \,;\Gamma \vdash M : A$. We show the two most difficult cases, namely the elimination rule, and that of $\mathsf{K4}$.

Case($\Delta \,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow P\ \mathsf{in}\ Q : C$).
By the IH, the premises imply that $\Delta^{\maltese}, \Gamma^{\maltese} \vdash P^{\maltese} : \mathbf{1} \times A^{\maltese}$, and $\Delta^{\maltese}, u : A^{\maltese}, \Gamma^{\maltese} \vdash Q^{\maltese} : C^{\maltese}$. Applying exchange multiple times, we obtain $\Delta^{\maltese}, \Gamma^{\maltese}, u : A^{\maltese} \vdash Q^{\maltese} : C^{\maltese}$, and using cut yields $\Delta^{\maltese}, \Gamma^{\maltese} \vdash Q^{\maltese}[\pi_2(P^{\maltese})/u] : C^{\maltese}$.

Case($\Delta \,;\Gamma \vdash_{\mathsf{DK4}} \mathsf{box}\ M : \Box A$).
By the IH, we have $\Delta^{\maltese}, \left(\Delta^{\perp}\right)^{\maltese} \vdash \left(M^{\perp}\right)^{\maltese} : A^{\maltese}$. Notice that, when acting on contexts, $(-)^{\perp}$ only acts on variables and $(-)^{\maltese}$ only on types, so $\left(\Delta^{\perp}\right)^{\maltese} \equiv \left(\Delta^{\maltese}\right)^{\perp}$. Substitute all the variables in $\left(\Delta^{\maltese}\right)^{\perp}$ with the corresponding ones in $\Delta^{\maltese}$. We thus obtain a term $\Delta^{\maltese} \vdash (M^{\perp})^{\maltese}[\Delta^{\maltese}/\left(\Delta^{\maltese}\right)^{\perp}] : A^{\maltese}$. But we have by (1) that

$$(M^{\perp})^{\maltese}[\Delta^{\maltese}/(\Delta^{\maltese})^{\perp}] \equiv (M^{\perp}[\Delta/(\Delta^{\maltese})^{\perp}])^{\maltese} \equiv (M^{\perp}[\Delta/\Delta^{\perp}])^{\maltese} \equiv M^{\maltese}$$

We thus apply weakening and the introduction rules for $*$ and products to obtain $\Delta^{\maltese}, \Gamma^{\maltese} \vdash \langle *, M^{\maltese}\rangle : \mathbf{1} \times A^{\maltese}$.

(3) By induction on $M \longrightarrow N$. The only cases that are not immediate are those involving modal constructors. For the modal β, we notice that by the congruence rules

$$(\mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ M\ \mathsf{in}\ N)^{\maltese} \equiv N^{\maltese}[\pi_2(\langle *, M^{\maltese}\rangle)/u] \longrightarrow N^{\maltese}[M^{\maltese}/x] \equiv (N[M/x])^{\maltese}$$

where the last step follows by (1). For the first commuting conversion, we have

$$(\pi_i(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N))^{\maltese} \equiv \pi_i(N^{\maltese}[\pi_2(M^{\maltese})/u]) \equiv (\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ \pi_i(N))^{\maltese}$$

so we use the fact $\longrightarrow^*_\beta$ is reflexive. The cases for the other commuting conversions are similar, and they translate to syntactic equalities. □

We can now almost obtain Theorem 4.9 as a corollary: an infinite sequence of reductions $N_1 \longrightarrow N_2 \longrightarrow \dots$ would lead to an infinite sequence of reductions $N_1^{\maltese} \longrightarrow^*_\beta N_2^{\maltese} \longrightarrow^*_\beta \dots$ in the simply-typed λ-calculus, which would contradict strong normalization. This is not yet a proof, because there is no guarantee that each reduction $N_i^{\maltese} \longrightarrow^*_\beta N_{i+1}^{\maltese}$ is a non-trivial β-reduction. However, scrutinising the proof of Theorem 4.10 leads to the conclusion that the only trivial reductions arise when $N_i \longrightarrow N_{i+1}$ due to a commuting conversion. It thus remains to prove that $N_i \longrightarrow N_{i+1}$ is infinitely often a proper β-reduction.

For that, we employ a trick used by de Groote [dG02]: we assign a *permutation degree* to every term. We define it by the following clauses.

$$|x| \stackrel{\text{def}}{=} 1 \qquad |\lambda x : A.\ M| \stackrel{\text{def}}{=} |M|$$

$$|MN| \stackrel{\text{def}}{=} |M| + \#(M) \cdot |N| \qquad |\langle M, N\rangle| \stackrel{\text{def}}{=} |M| + |N|$$

$$|\pi_i(M)| \stackrel{\text{def}}{=} |M| + \#(M) \qquad |\mathsf{box}\ M| \stackrel{\text{def}}{=} |M|$$

$$|\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N| \stackrel{\text{def}}{=} |M| + \#(M) \cdot |N|$$

where

$$\#(x) \stackrel{\text{def}}{=} 1 \qquad \#(\lambda x : A.\ M) \stackrel{\text{def}}{=} 1$$

$$\#(MN) \stackrel{\text{def}}{=} \#(M) \qquad \#(\langle M, N \rangle) \stackrel{\text{def}}{=} 1$$

$$\#(\pi_i(M)) \stackrel{\text{def}}{=} \#(M) \qquad \#(\mathsf{box}\ M) \stackrel{\text{def}}{=} 1$$

$$\#(\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N) \stackrel{\text{def}}{=} 2 \cdot \#(M) \cdot \#(N)$$

Briefly, $\#(M) = 2^n$, where n is the number of $\mathsf{let\ box}\ u \Leftarrow (-)\ \mathsf{in}\ (-)$ constructs in M that are either at an 'outermost' position, or can be commuted to be so. It follows that

Lemma 4.11. *If $M \longrightarrow N$ by a commuting conversion, then $\#(M) = \#(N)$.*

The metric $|M|$ uses $\#(M)$ to weigh the appearance of $\mathsf{let\ box}\ u \Leftarrow (-)\ \mathsf{in}\ (-)$ constructs, with the weight being higher the more deeply they appear in a term. It is easy to show that

Lemma 4.12. *If $M \longrightarrow N$ by a commuting conversion, then $|M| > |N|$.*

It follows that in an infinite sequence of reductions it must be the case that β-reductions occur infinitely often, as the metric $| - |$ strictly reduces for each commuting conversion. This completes the proof of strong normalization.

4.5. **The case of GL.** It remains to prove normalization for GL, which—because of the presence of a certain amount of self-reference—behaves in an unusual way.

Recall that $\mathsf{fix}\ z\ \mathsf{in\ box}\ M$ is the term corresponding to an application of the Löb rule to the proof M. One might first think that this term must reduce in the manner of a fixpoint, for example to $M[\mathsf{fix}\ z\ \mathsf{in\ box}\ M/z]$ where the diagonal variable $z : \Box A$ has been replaced by the entire term. Notice, however, that this does not preserve subject reduction. Instead, we follow the statement of modal cut admissibility (Theorem 3.8), and replace the modal β-reduction used in the other systems by

$$\mathsf{let\ box}\ u \Leftarrow (\mathsf{fix}\ z\ \mathsf{in\ box}\ M)\ \mathsf{in}\ N \longrightarrow N[M[\mathsf{fix}\ z\ \mathsf{in\ box}\ M/z]/u]$$

which preserves subject reduction. We also include the congruence rule

$$\frac{M \longrightarrow N}{\mathsf{fix}\ z\ \mathsf{in\ box}\ M \longrightarrow \mathsf{fix}\ z\ \mathsf{in\ box}\ N}$$

This leads to to a system with manifest *coinductive behaviour*: The introduction form $\mathsf{fix}\ z\ \mathsf{in\ box}\ M$ is a largely inactive term former which—contrary to expectation—does not unfold infinitely. Instead, we are free to use the congruence rule to eliminate cuts in its body M. On the other hand, when this term meets the elimination form $\mathsf{let\ box}\ u \Leftarrow (-)\ \mathsf{in}\ N$, it unfolds as much as necessary to fill the position of the variable u in N. Thus, $\mathsf{fix}\ z\ \mathsf{in\ box}\ M$ unfolds 'on demand' whenever it is deconstructed.

We can extend complement reduction (Lemma 4.1) to this system: if $\Delta\ ;\ \Delta^\perp, q^\perp : \Box A \vdash_{\mathsf{DGL}} M^\perp : A$, then $M \longrightarrow N$ implies $M^\perp \longrightarrow N^\perp$. The proof is the essentially the same, even in the case of the the new β-reduction: we have

$$(\mathsf{let\ box}\ u \Leftarrow \mathsf{fix}\ z\ \mathsf{in\ box}\ Q\ \mathsf{in}\ N)^\perp \longrightarrow N^\perp[Q_*/u]$$

where $Q_* \stackrel{\text{def}}{=} Q[\mathsf{fix}\ z\ \mathsf{in\ box}\ Q/z]$. We want to show that the RHS is $(N[Q_*/u])^\perp$. It suffices to show that $u^\perp \not\in \mathrm{Fv}\,(N)$ and $u^\perp \not\in \mathrm{Fv}\left(N^\perp\right)$, for then by Theorem 3.2 we get $(N[Q_*/u])^\perp \equiv$

$N^{\perp}[Q_*, Q_*^{\perp}/u, u^{\perp}] \equiv N^{\perp}[Q_*/u]$. Recalling that $\Delta, u^{\perp} : B \,;\, \Delta^{\perp}, q^{\perp} : \Box A \vdash N^{\perp} : A$ for some B, we see that the rest of the argument is essentially identical to that for $\mathsf{K4}$.

It is then easy to extend the proof of confluence and the subformula property to this notion of reduction. However, the method used to prove strong normalization in §4.4 no longer applies. Intuitively the reason is that, up to the isomorphism $A \cong \mathbf{1} \times A$, the Gödel-Löb axiom would be translated to the type $(A \to A) \to A$, which is the type of fixed point combinators at A. We would thus have to augment the target language with such fixed point combinators. This would make it essentially equivalent to PCF [Plo77], which is the archetypal language with non-terminating terms!

Instead, we will prove normalization using the method of candidates (candidats de reducibilité). This method originates in Girard's proof of strong normalization for System F [Gir72]. Our variant is a combination of two presentations. The main structure of the proof is due to by Koletsos [Kol85], as presented in simplified form by Gallier [Gal95]. However, the Koletsos-Gallier presentation does not carry typing information in the proof, whereas in our calculi typing is vital. Thus, we enhance their method, insofar as our can candidates consist of typing judgments $\Delta \,;\, \Gamma \vdash M : A$ rather than simply terms $M : A$. Ideas on how this is done were drawn from another chapter by Gallier [Gal90], which surveys multiple variants of the *candidats* method. The proof itself is rather long. We abbreviate it by only presenting the cases that are relevant to the modal fragment of the language. The interested reader can obtain the full proof from the author's website.[7]

The overall structure of the proof is the following. Suppose we have a family of nonempty *sets of typing judgments*,

$$\mathcal{P} = \{P_A\}_A$$

indexed by the type A they assign to the term they carry. If $C \subseteq P_A$, we write $\Delta ; \Gamma \vdash M \in C$ as a shorthand for $(\Delta \,;\, \Gamma \vdash M : A) \in C$. We will state six properties, $\mathsf{P0}$–$\mathsf{P5}$, that such a family should satisfy. In case it does indeed satisfy them, we show that

$$\Delta \,;\, \Gamma \vdash_{\mathsf{DGL}} M : A \implies \Delta \,;\, \Gamma \vdash M \in P_A$$

In our case, we show that the family of typing judgments

$$\mathcal{SN} \stackrel{\text{def}}{=} \{SN_A\}_A$$

satisfies the properties $\mathsf{P0}$-$\mathsf{P5}$, where SN_A consists of all the judgments $\Delta \,;\, \Gamma \vdash M : A$ for which M is strongly normalizing. Then $SN_A = \Lambda_A$, and all typable terms are strongly normalizing.

Here is a brief summary of the proof. We begin by stating the first four properties, namely $\mathsf{P0}$–$\mathsf{P3}$. We also define what it means for a set C of derivable judgments to be a *candidate*. Then, we define a subset $[\![A]\!] \subseteq P_A$ for each type A. We call judgments in $[\![A]\!]$ *reducible*, and we show $[\![A]\!]$ to be a candidate. Finally, we introduce two further properties, $\mathsf{P4}$ and $\mathsf{P5}$. If these hold of P_A, then we show that $[\![A]\!]$ contains all derivable judgments.

We write $\Gamma \sqsubseteq \Gamma'$ to mean that the context Γ is a subsequence of Γ' (in other words, Γ' is obtained after a series of weakening steps on Γ). We also write Λ for the set of all terms. We will make use of the following helpful definitions.

Definition 4.13.

(1) A term is a *intro term* just if it is an introduction form, i.e. of the form

$$\lambda x : A.\ M, \quad \langle M, N \rangle, \quad \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M$$

[7] `https://www.lambdabetaeta.eu`

(2) A term is a *simple term* just if it is a variable or an elimination form, i.e. of the forms

$$x, \quad MN, \quad \pi_i(M), \quad \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N$$

(3) A *stubborn term* is a term that is either a normal form w.r.t. $\longrightarrow$, or a term that does not reduce to an intro term, i.e. if $M \longrightarrow^* N$ then N is not an intro term.

Candidates and the first four properties. The sets of judgments we will use will always satisfy the following four important properties.

Definition 4.14 (Properties P0-P3)**.** For a family $\mathcal{P}$ we define the following properties.

(P0) (a) $\Delta\,;\Gamma \vdash M \in P_A$ and $\Gamma \sqsubseteq \Gamma'$ imply $\Delta\,;\Gamma' \vdash M \in P_A$
  (b) $\Delta\,;\Gamma \vdash M \in P_A$ and $\Delta \sqsubseteq \Delta'$ imply $\Delta'\,;\Gamma \vdash M \in P_A$

(P1) $\Delta\,;\Gamma \vdash x \in P_A$ for all $(x : A) \in \Gamma$.

(P2) $\Delta\,;\Gamma \vdash M \in P_A$ and $M \longrightarrow N$ imply $\Delta\,;\Gamma \vdash N \in P_A$.

(P3) For simple terms M,
  (a) If
  - $\Delta\,;\Gamma \vdash M \in P_{A\to B}$,
  - $\Delta\,;\Gamma \vdash N \in P_A$, and
  - whenever $M \longrightarrow^* \lambda x : A.\ M'$ then $\Delta\,;\Gamma \vdash (\lambda x : A.\ M')N \in P_B$

  then this implies that $\Delta\,;\Gamma \vdash MN \in P_B$.
  (b) If
  - $\Delta\,;\Gamma \vdash M \in P_{A\times B}$, and
  - whenever $M \longrightarrow^* \langle M_1, M_2\rangle$ then $\Delta\,;\Gamma \vdash \pi_1(\langle M_1, M_2\rangle) \in P_A$ and $\Delta\,;\Gamma \vdash \pi_2(\langle M_1, M_2\rangle) \in P_B$,

  then this implies that $\Delta\,;\Gamma \vdash \pi_1(M) \in P_A$ and $\Delta\,;\Gamma \vdash \pi_2(M) \in P_B$.

Definition 4.15 ($\mathcal{P}$-candidate)**.** A set $C_A \subseteq P_A$ is *$\mathcal{P}$-candidate at A* just if

(R0) (a) $\Delta\,;\Gamma \vdash M \in C_A$ and $\Gamma \sqsubseteq \Gamma'$ imply $\Delta\,;\Gamma' \vdash M \in C_A$.
  (b) $\Delta\,;\Gamma \vdash M \in C_A$ and $\Delta \sqsubseteq \Delta'$ imply $\Delta'\,;\Gamma \vdash M \in C_A$.

(R1) $\Delta\,;\Gamma \vdash M \in C_A$ and $M \longrightarrow N$ imply $\Delta\,;\Gamma \vdash N \in C_A$.

(R2) If $\Delta\,;\Gamma \vdash M \in P_A$ is simple, and $M \longrightarrow^* N$ and for an intro term N implies $\Delta\,;\Gamma \vdash N \in C_A$, then it follows that $\Delta\,;\Gamma \vdash M \in C_A$.

Notice that (R0) is analogous to (P0), and (R1) is analogous to (P2). Moreover, these conditions in tandem imply an analogue of (P1):

Lemma 4.16. *For any $\mathcal{P}$-candidate C_A, if $(x : A) \in \Gamma$ then $\Delta\,;\Gamma \vdash x \in C$.*

Proof. By (P1), we have that $\Delta\,;\Gamma \vdash x \in P_A$, and by definition x is simple, and a normal form, so it cannot ever reduce to an intro term. The result follows by (R2). □

A family $\mathcal{P}$ for which P0–P3 hold is almost a candidate. In fact, the only condition that is not automatically satisfied is R2. To remedy that situation, we define a particular subfamily $[\![A]\!] \subseteq P_A$ of *reducible judgments*, which—as we show—satisfies it. This definition has the familiar flavour of logical predicates.

Definition 4.17 (Reducible judgments)**.** We define for each type A a set of derivable judgments $[\![A]\!] \subseteq P_A$ by induction on A.

$$[\![p_i]\!] \stackrel{\text{def}}{=} P_{p_i}$$

$$[\![A \times B]\!] \stackrel{\text{def}}{=} \{\, \Delta\,;\Gamma \vdash M \in P_{A\times B} \mid \Delta\,;\Gamma \vdash \pi_1(M) \in [\![A]\!] \wedge\ \Delta\,;\Gamma \vdash \pi_2(M) \in [\![B]\!] \,\}$$

$$[\![A \to B]\!] \stackrel{\text{def}}{=} \left\{\, \Delta\,;\Gamma \vdash M \in P_{A\to B} \,\middle|\, \forall \Delta \sqsubseteq \Delta',\ \Gamma \sqsubseteq \Gamma', \Delta'\,;\Gamma' \vdash N \in [\![A]\!]\,.\ \Delta'\,;\Gamma' \vdash MN \in [\![B]\!] \,\right\}$$

$$[\![\Box A]\!] \stackrel{\text{def}}{=} \left\{\, \Delta\,;\Gamma \vdash M \in P_{\Box A} \,\middle|\, M \longrightarrow^* \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q \implies \Delta\,;\Delta^\perp, z^\perp : \Box A \vdash Q^\perp \in [\![A]\!] \,\right\}$$

We can now show that

Theorem 4.18. *If $\mathcal{P} = \{P_A\}$ satisfies properties* P0*-*P3*, then*
(1) *For any A, $[\![A]\!]$ is a $\mathcal{P}$-candidate.*
(2) *For any A, $[\![A]\!]$ contains all the stubborn terms in P_A.*

Proof. By induction on types.

CASE($\Box A$). For (1):

(R0) (a) trivially holds, for none of the judgments for Q in $[\![\Box A]\!]$ depend on Γ.
For (b): we first use (P0) to ascertain that $\Delta'\,;\Gamma \vdash M \in P_{\Box A}$. If M reduces to $\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q$, we use (R0)(a, b) of the IH for A to weaken Δ and $\Delta^\perp$ respectively.

(R1) Let $\Delta;\Gamma \vdash M \in [\![\Box A]\!]$ and suppose $M \longrightarrow N$. By (P2) we have $\Delta;\Gamma \vdash N \in P_{\Box A}$. It remains to show that whenever $N \longrightarrow^* \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q$ then $\Delta\,;\Delta'^\perp, z^\perp : \Box A \vdash Q^\perp \in [\![A]\!]$. But then $M \longrightarrow^* \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q$ as well, so this follows from $M \in [\![\Box A]\!]$.

(R2) Suppose that $\Delta\,;\Gamma \vdash M \in P_{\Box A}$ is a simple term, and whenever $M \longrightarrow^* \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q$ then $\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q \in [\![\Box A]\!]$, i.e. $\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q \longrightarrow^* \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q'$ implies $\Delta\,;\Delta^\perp, z^\perp : \Box A \vdash Q'^\perp \in [\![A]\!]$. We need to show that, if $M \longrightarrow^* \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q$, then $\Delta\,;\Delta^\perp, z^\perp : \Box A \vdash Q \in [\![A]\!]$. But this follows by the reflexivity of $\longrightarrow^*$.

For (2): if $M \in P_{\Box A}$ is stubborn, then it never reduces to an intro term $\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q$, so it is vacuously in $[\![\Box A]\!]$. ∎

Closure under formation: the latter two properties. Unfortunately, this is not enough to show that the candidates $[\![A]\!]$ contain *all* the provable judgments of DGL. We will thus need the following two additional conditions on $\mathcal{P}$.

Definition 4.19 (Properties P4-P5)**.**

(P4) (a) If $\Delta\,;\Gamma, x : A \vdash M \in P_B$ then $\Delta\,;\Gamma \vdash \lambda x : A.\ M \in P_{A\to B}$.
(b) $\Delta\,;\Gamma \vdash M \in P_A$ and $\Delta\,;\Gamma \vdash N \in P_B$ imply $\Delta\,;\Gamma \vdash \langle M, N\rangle \in P_{A\times B}$.
(c) $\Delta\,;\Delta^\perp, z^\perp : \Box A \vdash Q^\perp \in P_A$ implies $\Delta\,;\Gamma \vdash \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q \in P_{\Box A}$

(P5) (a) $\Delta \sqsubseteq \Delta'$, $\Gamma \sqsubseteq \Gamma'$, $\Delta'\,;\Gamma' \vdash N \in P_A$, and $\Delta'\,;\Gamma' \vdash M[N/x] \in P_B$ imply $\Delta'\,;\Gamma' \vdash (\lambda x : A.\ M)N \in P_B$.
(b) $\Delta\,;\Gamma \vdash M \in P_A$ and $\Delta\,;\Gamma \vdash N \in P_B$ imply $\Delta\,;\Gamma \vdash \pi_1(\langle M, N\rangle) \in P_A$ and $\Delta\,;\Gamma \vdash \pi_2(\langle M, N\rangle) \in P_B$.
(c) If $\Delta;\Gamma \vdash M \in P_{\Box A}$ and $\Delta, u : A;\Gamma \vdash N \in P_C$, and whenever $M \longrightarrow^* \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q$ then $\Delta\,;\Gamma \vdash N\,[Q[\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q/z]/u] \in P_C$, then $\Delta\,;\Gamma \vdash \mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N \in P_C$.

These conditions ensure that the candidates $[\![A]\!]$ also have the following closure properties.

Theorem 4.20. *If $\mathcal{P} = \{P_A\}$ satisfies properties P0–P5, then*

(1) *If whenever $\Gamma \sqsubseteq \Gamma'$, $\Delta \sqsubseteq \Delta'$ and $\Delta' ; \Gamma' \vdash N \in [\![A]\!]$ we have $\Delta' ; \Gamma' \vdash M[N/x] \in [\![B]\!]$, then $\Delta ; \Gamma \vdash \lambda x : A.\ M \in [\![A \to B]\!]$.*
(2) *If $\Delta ; \Gamma \vdash M \in [\![A]\!]$ and $\Delta ; \Gamma \vdash N \in [\![B]\!]$ then $\Delta ; \Gamma \vdash \langle M, N\rangle \in [\![A \times B]\!]$.*
(3) *If $\Delta ; \Gamma \vdash M \in [\![\Box A]\!]$, and whenever $\Delta \sqsubseteq \Delta'$ and $\Delta' ; \Delta'^{\perp}, z^{\perp} : \Box A \vdash Q^{\perp} \in [\![A]\!]$ then $\Delta' ; \Gamma \vdash N\,[Q[\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q/z]/u] \in [\![C]\!]$, then $\Delta ; \Gamma \vdash \mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N \in [\![C]\!]$.*

Proof. We only show (3). Write $Q_\star \stackrel{\text{def}}{=} Q[\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q/z]$.

First, we show that $\mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N \in P_C$, and we invoke (P5)(c) to do so. It suffices to show that $\Delta ; \Gamma \vdash M \in P_{\Box A}$, that $\Delta, u : A ; \Gamma \vdash N \in P_C$, and that whenever $M \longrightarrow^* \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q$ then $\Delta ; \Gamma \vdash N[Q_\star/u] \in P_C$. The first of these is implied by the assumption that $\Delta ; \Gamma \vdash M \in [\![A]\!] \subseteq P_A$. For the second, we infer by Lemma 4.16 and Theorem 4.18 that $\Delta, u : A ; \Delta^{\perp}, u^{\perp} : A, z^{\perp} : A \vdash u^{\perp} \in [\![A]\!]$. Hence, as $\Delta \sqsubseteq \Delta, u : A$, we have by the assumption that

$$\Delta, u : A ; \Gamma \vdash N \equiv N[u[\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ u/z]/u] \in [\![C]\!]$$

The final desideratum also follows: if $M \longrightarrow^* \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q$ then, by the definition of $[\![\Box A]\!]$, we have that $\Delta ; \Delta^{\perp}, z^{\perp} : \Box A \vdash Q^{\perp} \in [\![A]\!]$ and hence—by the assumption—that $\Delta ; \Gamma \vdash N[Q_\star/u] \in [\![C]\!] \subseteq P_C$.

For the rest, we note that $\mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N$ is simple, so we use (R2): it suffices to show that whenever $\mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N \longrightarrow^* Q$ and Q is an intro term, then $Q \in [\![C]\!]$. If $\mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N$ is stubborn, then the desideratum is trivial. Otherwise, if $\mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N \longrightarrow^* Q$ where Q is a intro term, then the reduction must be of the form

$$\begin{aligned} \mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N \longrightarrow^* &\quad \mathsf{let}\ \mathsf{box}\ u \Leftarrow \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ U\ \mathsf{in}\ N' \\ \longrightarrow &\quad N'[U^\star/u] \\ \longrightarrow^* &\quad Q \end{aligned}$$

where $M \longrightarrow^* \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ U$ and $N \longrightarrow^* N'$: otherwise the let construct would persist. But, by assumption, $\Delta ; \Gamma \vdash M \in [\![\Box A]\!]$, so by multiple applications of (R1) we infer that $\Delta;\Gamma \vdash \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ U \in [\![\Box A]\!]$ and hence that $\Delta;\Delta', z^{\perp} : \Box A \vdash U^{\perp} \in [\![A]\!]$. By the assumption, we get $\Delta ; \Gamma \vdash N[U_\star/u] \in [\![C]\!]$. But $N[U_\star/u] \longrightarrow^* N'[U_\star/u] \longrightarrow^* Q$, so $Q \in [\![C]\!]$ by repeated applications of (R1). ☐

The main theorem.

Definition 4.21 (Substitution)**.**

(1) A *substitution* is a finite function $\sigma : \mathcal{V} \rightharpoonup \Lambda$ from the set of all variables $\mathcal{V}$ to the set of all possible untyped/raw terms Λ.
(2) A substitution σ is *type-preserving* from $\Delta' ; \Gamma'$ to $\Delta ; \Gamma$, written $\Delta' ; \Gamma' \stackrel{\sigma}{\Rightarrow} \Delta ; \Gamma$, just if
   (a) $dom(\sigma) \subseteq \textsc{Vars}\,(\Delta) \cup \textsc{Vars}\,(\Gamma)$,
   (b) $(x : B) \in \Gamma$ implies $\Delta' ; \Gamma' \vdash \sigma(x) : B$, and
   (c) there exists z such that $(u : B) \in \Delta$ implies $\Delta' ; \Delta'^{\perp}, z^{\perp} : \Box B \vdash (\sigma(u))^{\perp} \in B$.

We call the aforementioned z the *diagonal variable* of the substitution. We have that

Lemma 4.22. *If $\Delta' ; \Gamma' \stackrel{\sigma}{\Rightarrow} \Delta ; \Gamma$ and $\Delta' \sqsubseteq \Delta''$ and $\Gamma' \sqsubseteq \Gamma'$ then $\Delta'' ; \Gamma'' \stackrel{\sigma}{\Rightarrow} \Delta ; \Gamma$.*

We write $\sigma[x \mapsto N]$ to mean the substitution defined by $\sigma[x \mapsto N](x) \stackrel{\text{def}}{=} N$, and $\sigma[x \mapsto N](y) \stackrel{\text{def}}{=} \sigma(y)$ if $y \not\equiv x$. Furthermore, if $\Delta' ; \Gamma' \stackrel{\sigma}{\Rightarrow} \Delta ; \Gamma$ with diagonal variable z, we define $\sigma^{\dagger}$ by

$$\sigma^{\dagger}(y) \stackrel{\text{def}}{=} \begin{cases} z^{\perp} & \text{if } y \equiv z^{\perp} \\ (\sigma(u))^{\perp} & \text{if } y \equiv u^{\perp} \in \textsc{Vars}\left(\Delta^{\perp}\right) \\ \sigma(y) & \text{otherwise} \end{cases}$$

The following technical fact is evident, but very convenient.

Lemma 4.23 (Modal Drop)**.** *If $\Delta' ; \Gamma' \stackrel{\sigma}{\Rightarrow} \Delta ; \Gamma$ with diagonal variable z, then*

$$\Delta' ; \Delta'^{\perp}, z^{\perp} : \Box A \stackrel{\sigma^{\dagger}}{\Longrightarrow} \Delta ; \Delta^{\perp}, z^{\perp} : \Box A$$

We extend the action of substitutions on terms in the usual capture-avoiding manner, e.g.

$$\sigma(\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M) \stackrel{\text{def}}{=} \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ \sigma(M)$$

Lemma 4.24. *If $\Delta' ; \Gamma' \stackrel{\sigma}{\Rightarrow} \Delta ; \Gamma$ and $\Delta ; \Gamma \vdash M : C$ then $\Delta' ; \Gamma' \vdash \sigma(M) : C$.*

Proof. By induction on M. We only show the modal cases. That of $\mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N$ is very similar to λ-abstraction. For $\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M$, we have that $\Delta ; \Delta^{\perp}, z^{\perp} : \Box A \vdash M : A$ for some A such that $C \equiv \Box A$. It follows by Lemma 4.23 that

$$\Delta' ; \Delta'^{\perp}, z^{\perp} : \Box A \stackrel{\sigma^{\dagger}}{\Longrightarrow} \Delta ; \Delta^{\perp}, z^{\perp} : \Box A$$

Applying the IH then yields $\Delta' ; \Delta'^{\perp}, z^{\perp} : \Box A \vdash \sigma^{\dagger}(M) : A$. But notice that $\sigma^{\dagger}(M) \equiv (\sigma(M))^{\perp}$, so a single use of $(\Box\mathcal{I}_{\mathsf{GL}})$ suffices. $\Box$

Theorem 4.25 (Candidats)**.** *Let $\mathcal{P} = \{P_A\}$ be a family satisfying properties P1–P5. Let $\Delta ; \Gamma \vdash_{\mathsf{DGL}} M : A$, and $\Delta' ; \Gamma' \stackrel{\sigma}{\Rightarrow} \Delta ; \Gamma$ be a substitution with diagonal variable z that respects the candidates, i.e. such that*

(1) $(x : B) \in \Gamma$ *implies* $\Delta' ; \Gamma' \vdash \sigma(x) \in [\![B]\!]$, *and*
(2) $(u : C) \in \Delta$ *implies* $\Delta' ; \Delta'^{\perp}, z^{\perp} : \Box C \vdash (\sigma(u))^{\perp} \in [\![C]\!]$,

Then $\Delta' ; \Gamma' \vdash \sigma(M) \in [\![A]\!]$.

Proof. By induction on M. We only prove the modal cases.

Case($\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M$).

Then $\Delta ; \Delta^{\perp}, z^{\perp} : \Box A \vdash M : A$. By Lemma 4.23, we have that $\Delta' ; \Delta'^{\perp}, z^{\perp} : \Box A \stackrel{\sigma^{\dagger}}{\Longrightarrow} \Delta ; \Delta^{\perp}, z^{\perp} : \Box A$. Then, by the IH we have that $\Delta' ; \Delta'^{\perp}, z^{\perp} : \Box A \vdash \sigma^{\dagger}(M) \equiv \sigma(M)^{\perp} \in [\![A]\!]$. So, by (P4)(c), $\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ \sigma(M) \in P_{\Box A}$. It now suffices—by the definition of $[\![\Box A]\!]$—to show that

$$\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ \sigma(M) \longrightarrow^{*} \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M'$$

implies $\Delta' ; \Delta'^{\perp}, z^{\perp} : \Box A \vdash M' \in [\![A]\!]$. But then we must have $\sigma(M) \longrightarrow^{*} M'$, so by repeated applications of (R1) we have $M' \in [\![A]\!]$.

Case($\mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N$).

We show the case for K. We have $\Delta ; \Gamma \vdash M : A$ and $\Delta, u : A ; \Gamma \vdash N : C$. We use Theorem 4.20(a): to show that

$$\Delta' ; \Gamma' \vdash \sigma(\mathsf{let}\ \mathsf{box}\ u \Leftarrow M\ \mathsf{in}\ N) \equiv \mathsf{let}\ \mathsf{box}\ u \Leftarrow \sigma(M)\ \mathsf{in}\ \sigma(N) \in [\![C]\!]$$

It suffices to show that $\Delta' ; \Gamma' \vdash \sigma(M) \in [\![\Box A]\!]$—which we have by the IH—and that whenever $\Delta'' \sqsupseteq \Delta'$ and $\cdot ; \Delta'' \vdash Q \in [\![A]\!]$, then $\Delta'' ; \Gamma' \vdash \sigma(N)[Q/u] \in [\![C]\!]$.

Define

$$\sigma' \stackrel{\text{def}}{=} \sigma[u \mapsto Q]$$

Then, by weakening the modal context in σ, we have

$$\Delta'', u : A ; \Gamma' \stackrel{\sigma'}{\Rightarrow} \Delta, u : A ; \Gamma$$

By the IH,

$$\Delta'', u : A ; \Gamma' \vdash \sigma'(N) \in [\![C]\!]$$

But $\sigma'(N) \equiv \sigma(N)[Q/u]$. □

Corollary 4.26. *If $\mathcal{P} = \{P_A\}$ is a family satisfying properties P0–P5, then*

$$P_A = \Lambda_A$$

Proof. By Theorem 4.25 we have that $\Delta ; \Gamma \vdash M \in [\![A]\!]$ for every $\Delta ; \Gamma \vdash M : A$. Hence $\Lambda_A \subseteq [\![A]\!] \subseteq P_A \subseteq \Lambda_A$. □

It is then reasonably straightforward to verify that the above properties hold of the family $\mathcal{SN}$. In carrying out the proof we shall often use induction on $d(M)$, the *depth of the term* M. We argue that this admissible as follows. First, we construct the *reduction tree* of M, which consists of M and all its reducts, with an edge from reduct M_1 to reduct M_2 just if $M_1 \longrightarrow M_2$. As M has at most finite redexes, the reduction tree is finitely branching: there can only be a finite number of terms M_i such that $N \longrightarrow M_i$ for any term N. Furthermore, if M is strongly normalizing, then the reduction tree has no infinite paths. By König's Lemma, the tree is then finite, and $d(M)$ is the depth of the reduction tree of M—i.e. the longest path in the tree that is rooted at M. Because of this use of König's Lemma, it is not clear whether this proof is constructive. We once more only prove the modal cases.

(P4)(c): If $\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q \longrightarrow P$, then $P \equiv \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ Q'$ for some Q' with $Q \longrightarrow Q'$. Hence $d(\mathsf{box}\ M) \leq d(M)$. But the last one is less than or equal to $d(M^{\perp})$ by Lemma 4.1, which is finite by assumption.

(P5)(c): First, we note that by substituting u for Q, the premise implies that N is strongly normalizing, and thus that both $d(M)$ and $d(N)$ are finite.

We now proceed by induction on $d(M) + d(N)$. If $\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N \longrightarrow P$, then there are three possibilities:

- $P \equiv \mathsf{let\ box}\ u \Leftarrow M'\ \mathsf{in}\ N$ and $M \longrightarrow M'$. Then

$$d(M') + d(N) < d(M) + d(N)$$

  and so, by the IH, P is strongly normalizing.
- Likewise for N.
- $M \equiv \mathsf{box}\ Q$ and $P \equiv N[Q/u]$. Then, by assumption, P is strongly normalizing.

In all cases, if $\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N \longrightarrow P$, then P is strongly normalizing. We conclude that the original term itself is strongly normalizing.

5. Modal category theory

We will now introduce a modest amount of monoidal category theory that is necessary for the formulation of a categorical semantics. This is needed because we will model the modality by a *strong monoidal endofunctor*. In our case the monoidal product will always be the cartesian product of a cartesian closed category. We will show that this coincides with the notion of *product-preserving endofunctor*, and hence gives rise to the isomorphism $\Box(A \times B) \cong \Box A \times \Box B$, which is another way of stating the modal axiom K (§1.3).

We assume some familiarity with the relationship between λ-calculi and cartesian closed categories (CCCs). A category $\mathcal{C}$ with finite products is a CCC whenever for each pair $A, B \in \mathcal{C}$ there is an object $B^A \in \mathcal{C}$ and a morphism $\mathsf{ev}_{A,B} : B^A \times A \to B$ such that for every $f : C \times A \to B$ there is a unique $\lambda(f) : C \to B^A$ such that $\mathsf{ev}_{A,B} \circ (\lambda(f) \times id_A) = f$. For further background on CCCs and the typed λ-calculus, we refer the reader to [LS88, Cro93, Awo10, AT11]. Some material on monoidal category theory is drawn from the superbly lucid treatment by Melliès [Mel09, §5], which is specifically geared towards categorical logic. Further material can be found in [Mac78, §XI.2].

5.1. Lax and strong monoidal functors.

Let $\mathcal{C}$ and $\mathcal{D}$ be cartesian categories. We regard them as monoidal categories $(\mathcal{C}, \times, \mathbf{1})$ and $(\mathcal{D}, \times, \mathbf{1})$ respectively.

Definition 5.1. A functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ between two cartesian categories is *lax monoidal* just if it is equipped with a natural transformation $m : F(-) \times F(-) \Rightarrow F(- \times -)$ as well as an arrow $m_0 : \mathbf{1} \to F(\mathbf{1})$ such that the following diagrams commute:

$$\begin{tikzcd}
(FA \times FB) \times FC \arrow[r, "\alpha"] \arrow[d, "m_{A,B} \times id_{FC}"'] & FA \times (FB \times FC) \arrow[d, "id_{FA} \times m_{B,C}"] \\
F(A \times B) \times FC \arrow[d, "m_{A \times B, C}"'] & FA \times F(B \times C) \arrow[d, "m_{A, B \times C}"] \\
F((A \times B) \times C) \arrow[r, "F\alpha"'] & F(A \times (B \times C))
\end{tikzcd} \tag{5.1}$$

$$\begin{tikzcd}
FA \times \mathbf{1} \arrow[r, "\rho_A"] \arrow[d, "id_{FA} \times m_0"'] & FA & \mathbf{1} \times FB \arrow[r, "\lambda_B"] \arrow[d, "m_0 \times id_{FB}"'] & FB \\
FA \times F\mathbf{1} \arrow[r, "m_{A,\mathbf{1}}"'] & F(A \times \mathbf{1}) \arrow[u, "F\rho_A"'] & F\mathbf{1} \times FB \arrow[r, "m_{\mathbf{1},B}"'] & F(\mathbf{1} \times B) \arrow[u, "F\lambda_B"']
\end{tikzcd} \tag{5.2}$$

Definition 5.2. A *strong monoidal functor* between two cartesian categories is a lax monoidal functor where the components $m_{A,B} : FA \times FB \to F(A \times B)$ and the arrow $m_0 : \mathbf{1} \to F\mathbf{1}$ are isomorphisms.

These natural transformations can be extended to arbitrary contexts. We write

$$\prod_{i=1}^{n} A_n \stackrel{\text{def}}{=} A_1 \times \cdots \times A_n$$

where $\times$ associates to the left. Then, we define the following morphisms by induction:

$$m^{(0)} \stackrel{\text{def}}{=} \mathbf{1} \xrightarrow{m_0} F\mathbf{1}$$

$$m^{(1)} \stackrel{\text{def}}{=} FA_1 \xrightarrow{id_{FA_1}} FA_1$$

$$m^{(n+1)} \stackrel{\text{def}}{=} \prod_{i=1}^{n+1} FA_i \xrightarrow{m^{(n)} \times id} F\left(\prod_{i=1}^{n} A_i\right) \times FA_{n+1} \xrightarrow{m} F\left(\prod_{i=1}^{n+1} A_i\right)$$

Then the $m^{(n)} : \prod_{i=1}^{n} FA_n \to F\left(\prod_{i=1}^{n} A_i\right)$'s are a natural transformation, i.e.

$$m^{(n)} \circ \prod_{i=1}^{n} Ff_i = F\left(\prod_{i=1}^{n} f_i\right) \circ m^{(n)}$$

We also note that if $F : \mathcal{C} \longrightarrow \mathcal{C}$ is a monoidal endofunctor, then so is $F^2 \stackrel{\text{def}}{=} F \circ F$, with

$$n_{A,B} \stackrel{\text{def}}{=} F^2A \times F^2B \xrightarrow{m_{A,B}} F(FA \times FB) \xrightarrow{Fm_{A,B}} F^2(A \times B)$$

and $n_0 \stackrel{\text{def}}{=} Fm_0 \circ m_0$. See [Mel09, §5.9].

5.2. **Product-Preserving Functors.** Lax and strong monoidal functors are widely used as morphisms between monoidal categories. However, our monoidal product will always be the cartesian product, so it is worth examining how these notions adapt to this particular setting. To begin, we should compare them to another kind of morphism between cartesian categories that 'plays well with products,' namely that of *product-preserving functors.*

Definition 5.3. A *product-preserving functor* $F : \mathcal{C} \longrightarrow \mathcal{D}$ between two cartesian categories is a functor for which the canonical arrows

$$p_{A,B} \stackrel{\text{def}}{=} \langle F\pi_1, F\pi_2 \rangle : F(A \times B) \xrightarrow{\cong} FA \times FB \qquad\qquad p_0 \stackrel{\text{def}}{=} !_{F\mathbf{1}} : F\mathbf{1} \xrightarrow{\cong} \mathbf{1}$$

are isomorphisms.

This definition appears to be much stronger. Indeed, product-preserving functors are strong monoidal. To show that, all we need to do is show that the inverses

$$m_{A,B} \stackrel{\text{def}}{=} p_{A,B}^{-1} : FA \times FB \xrightarrow{\cong} F(A \times B) \qquad\qquad m_0 \stackrel{\text{def}}{=} p_0^{-1} : \mathbf{1} \xrightarrow{\cong} F\mathbf{1}$$

satisfy the coherence conditions (5.1) and (5.2). Before we do that, we note two very useful equations that product-preserving functors satisfy. The first is

Proposition 5.4. *If F is product-preserving then $m_{A,B} \circ \langle Ff, Fg \rangle = F\langle f, g \rangle$.*

Proof. Calculate that $p_{A,B} \circ F\langle f, g \rangle = \langle Ff, Fg \rangle$ and notice $p_{A,B}^{-1} = m_{A,B}$. $\square$

The second equation shows how $m_{A,B}$ may be used to relate projections.

Proposition 5.5. *Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be product-preserving, and let $A \xleftarrow{\pi_1^{A,B}} A \times B \xrightarrow{\pi_2^{A,B}} B$ and $FA \xleftarrow{\pi_1^{FA,FB}} FA \times FB \xrightarrow{\pi_2^{FA,FB}} FB$ be product diagrams in $\mathcal{C}$ and $\mathcal{D}$ respectively. Then*

$$F\pi_i^{A,B} \circ m_{A,B} = \pi_i^{FA,FB}$$

Proof. Calculate that $\pi_i \circ m_{A,B}^{-1} = \pi_i \circ \langle F\pi_1, F\pi_2 \rangle = F\pi_i$ and notice $p_{A,B}^{-1} = m_{A,B}$. $\square$

We will often write such equations as $F\pi_1 \circ m = \pi_1$ without further ado. Armed with these facts, we see that the definitions of $m_{A,B}$ and m_0 given above are natural, and that

Theorem 5.6. *Any product-preserving functor is strong monoidal.*

Rather strikingly, the converse holds as well.

Theorem 5.7. *Any functor that is strong monoidal with respect to the cartesian structure is product-preserving.*

Proof. Note that $m_0^{-1} : F\mathbf{1} \longrightarrow \mathbf{1}$ is necessarily equal to the unique arrow $!_{F\mathbf{1}} : F\mathbf{1} \longrightarrow \mathbf{1}$ to the terminal object $\mathbf{1}$. Hence, it suffices to show that for any $A, B \in \mathcal{C}$, $m_{A,B}^{-1} = \langle F\pi_1, F\pi_2\rangle$. We will first show a particular case, viz. that

$$m_{A,\mathbf{1}}^{-1} = \langle F\pi_1, F\pi_2\rangle$$

from which the general case will follow. Remembering that $\rho_A \stackrel{\text{def}}{=} \pi_1 : A \times \mathbf{1} \to A$, we have that $\rho_A^{-1} = \langle id_{FA}, !_{FA}\rangle : FA \to FA \times \mathbf{1}$. Hence, by reversing ρ_A and $m_{A,\mathbf{1}}$ in the first diagram of (5.2) we obtain

$$m_{A,\mathbf{1}}^{-1} = (id_A \times m_0) \circ \langle id_{FA}, !_{FA}\rangle \circ F\pi_1 = \langle F\pi_1, m_0 \circ !_{F(A\times\mathbf{1})}\rangle : F(A \times \mathbf{1}) \to FA \times F\mathbf{1}$$

But, as $m_0 : \mathbf{1} \xrightarrow{\cong} F\mathbf{1}$, $F\mathbf{1}$ is also a terminal object, and any arrow into it is of the form $m_0 \circ !_A : A \to F\mathbf{1}$. This applies to $F\pi_2 : F(A \times \mathbf{1}) \to F\mathbf{1}$, so $m_{A,\mathbf{1}}^{-1} = \langle F\pi_1, F\pi_2\rangle$.

Now for the general case. As $m_{A,B}$ is a natural isomorphism, its inverse is a natural transformation with components $m_{A,B}^{-1}$. The naturality square for $(id_A, !_B)$ is

$$\begin{tikzcd}
F(A \times B) \arrow[r, "m_{A,B}^{-1}"] \arrow[d, "F(id_A \times !_B)"'] & FA \times FB \arrow[d, "id_{FA} \times F(!_B)"] \\
F(A \times \mathbf{1}) \arrow[r, "m_{A,\mathbf{1}}^{-1}"'] & FA \times F\mathbf{1}
\end{tikzcd}$$

Calculating down and across gives

$$m_{A,\mathbf{1}}^{-1} \circ F(id_A \times !_B) = \langle F\pi_1, F\pi_2\rangle \circ F(id_A \times !_B) = \langle F\pi_1, F(!_B \circ \pi_2)\rangle$$

whereas across and down gives

$$(id_{FA} \times F(!_B)) \circ m_{A,B}^{-1} = \langle \pi_1 \circ m_{A,B}^{-1}, F(!_B) \circ \pi_2 \circ m_{A,B}^{-1}\rangle$$

The first two components of these should be equal, therefore $\pi_1 \circ m_{A,B}^{-1} = F\pi_1$. Similarly, $\pi_2 \circ m_{A,B}^{-1} = F\pi_2$, and hence $m_{A,B}^{-1} = \langle F\pi_1, F\pi_2\rangle$. $\square$

5.3. **Monoidal natural transformations.** The standard notion of natural transformation between lax monoidal functors is the following.

Definition 5.8. Let $F, G : \mathcal{C} \longrightarrow \mathcal{D}$ be two lax monoidal functors between two cartesian categories. A *monoidal natural transformation* between F and G is a natural transformation

$\alpha : F \Rightarrow G$ such that the following diagrams commute:

$$\begin{tikzcd} FA \times FB \arrow[r, "\alpha_A \times \alpha_B"] \arrow[d, "m_{A,B}"'] & GA \times GB \arrow[d, "n_{A,B}"] \\ F(A \times B) \arrow[r, "\alpha_{A\times B}"'] & G(A \times B) \end{tikzcd} \qquad \begin{tikzcd} \mathbf{1} \arrow[d, "m_0"'] \arrow[dr, "n_0"] & \\ F\mathbf{1} \arrow[r, "\alpha_{\mathbf{1}}"'] & G\mathbf{1} \end{tikzcd}$$

Surprisingly, it is not hard to show that

Theorem 5.9. *If $F, G : \mathcal{C} \longrightarrow \mathcal{D}$ are product-preserving functors between two cartesian categories, then any natural transformation $\alpha : F \Rightarrow G$ is a monoidal natural transformation.*

Proof. We trivially have $!_{G\mathbf{1}} \circ \alpha_{\mathbf{1}} = !_{F\mathbf{1}}$. But $!_{G\mathbf{1}}$ and $!_{F\mathbf{1}}$ are isomorphisms, so—by inverting them—we obtain $\alpha_{\mathbf{1}} \circ m_0 = n_0$. Furthermore, we have the following naturality diagram:

$$\begin{tikzcd} F(A \times B) \arrow[r, "F\pi_1"] \arrow[d, "\alpha_{A\times B}"'] & FA \arrow[d, "\alpha_A"] \\ G(A \times B) \arrow[r, "G\pi_1"'] & GA \end{tikzcd}$$

and a similar one for B. Hence, $n^{-1}_{A,B} \circ \alpha_{A\times B} \stackrel{\text{def}}{=} \langle G\pi_1, G\pi_2 \rangle \circ \alpha_{A\times B}$ is equal to

$$\langle G\pi_1 \circ \alpha_{A\times B}, G\pi_2 \circ \alpha_{A\times B} \rangle = \langle \alpha_A \circ F\pi_1, \alpha_B \circ F\pi_2 \rangle = (\alpha_A \times \alpha_B) \circ \langle F\pi_1, F\pi_2 \rangle$$

which is just $(\alpha_A \times \alpha_B) \circ m^{-1}_{A,B}$. It then suffices to invert $m_{A,B}$ and $n_{A,B}$. □

5.4. **The categorical interpretation of modal rules.** In this section we introduce the main structures needed to produce categorical models for our calculi. We begin with the basic two examples of Kripke categories (K), and Bierman–de Paiva categories (S4). These are the most well-behaved, and most commonly encountered cases. We then discuss the slightly more obscure cases of Kripke-4 categories (K4), Kripke-T categories (T), and finally Gödel-Löb categories (GL).

5.4.1. *Kripke categories.* The combination of a CCC with a product-preserving endofunctor is the quintessential structure in our development, so we give it a name.

Definition 5.10. A *Kripke category* $(\mathcal{C}, \times, \mathbf{1}, F)$ is a cartesian closed category $\mathcal{C}$ along with a *product-preserving endofunctor* $F : \mathcal{C} \longrightarrow \mathcal{C}$.

Kripke categories are the minimal setting in which one can model Scott's rule (see §1.5.2), by defining an operation

$$(-)^\bullet : \mathcal{C}\left(\prod_{i=1}^n A_i, B\right) \to \mathcal{C}\left(\prod_{i=1}^n FA_i, FB\right)$$

by

$$f : \prod_{i=1}^n A_i \to B \quad \longmapsto \quad f^\bullet \stackrel{\text{def}}{=} \prod_{i=1}^n FA_i \xrightarrow{m^{(n)}} F\left(\prod_{i=1}^n A_i\right) \xrightarrow{Ff} FB$$

The operation $(-)^\bullet$ satisfies the following distribution/naturality laws.

Proposition 5.11.

(1) *Let* $f : \prod_{i=1}^{n} B_i \to C$ *and* $g_i : \prod_{j=1}^{k} A_j \to B_i$ *for* $i = 1, \ldots, n$. *Then*

$$(f \circ \langle \overrightarrow{g_i} \rangle)^{\bullet} = f^{\bullet} \circ \left\langle \overrightarrow{g_i^{\bullet}} \right\rangle$$

(2) *For* $f : \prod_{i=1}^{n} A_i \to B$ *and* $\langle \vec{\pi_j} \rangle : \prod_{i=1}^{n} FA_i \to \prod_{j \in J} FA_j$ *for* J *a list drawn from* $\{1, \ldots, n\}$,

$$(f \circ \langle \vec{\pi_j} \rangle)^{\bullet} = f^{\bullet} \circ \langle \vec{\pi_j} \rangle$$

As product-preserving endofunctors abound in the literature on category-theory, we do not see any use in giving examples of Kripke categories. Rather, their omnipresence is an attestation of the importance of our system as a candidate internal language.

5.4.2. *Bierman-de Paiva categories.* In order to model $\mathsf{S4}$, we need a Kripke category whose product-preserving functor is additionally a *comonad*.

Definition 5.12. A comonad (F, ϵ, δ) consists of an endofunctor $F : \mathcal{C} \longrightarrow \mathcal{C}$, and two natural transformations

$$\epsilon : F \Rightarrow \mathsf{Id}, \qquad \delta : F \Rightarrow F^2$$

such that the following diagrams commute:

$$\begin{tikzcd}
FA \arrow[r, "\delta_A"] \arrow[d, "\delta_A"'] & F^2A \arrow[d, "\delta_{FA}"] \\
F^2A \arrow[r, "F\delta_A"'] & F^3A
\end{tikzcd}
\qquad
\begin{tikzcd}
FA \arrow[r, "\delta_A"] \arrow[d, "\delta_A"'] \arrow[dr, "id_{FA}"] & F^2A \arrow[d, "\epsilon_{FA}"] \\
F^2A \arrow[r, "F\epsilon_A"'] & FA
\end{tikzcd}$$

In particular, we will require that the comonad used is *monoidal*, in that it satisfies some additional coherence conditions with respect to the monoidality.

Definition 5.13. A *monoidal comonad* on a cartesian category $\mathcal{C}$ is a comonad (F, ϵ, δ) such that $F : \mathcal{C} \longrightarrow \mathcal{C}$ is a lax monoidal functor, and $\epsilon : F \Rightarrow \mathsf{Id}$ and $\delta : F \Rightarrow F^2$ are monoidal natural transformations. Concretely, this amounts to the commutation of

$$\begin{tikzcd}
FA \times FB \arrow[r, "\epsilon_A \times \epsilon_B"] \arrow[d, "m_{A,B}"'] & A \times B \arrow[d, equal] \\
F(A \times B) \arrow[r, "\epsilon_{A \times B}"'] & A \times B
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{1} \arrow[d, "m_0"'] \arrow[dr, equal] & \\
F\mathbf{1} \arrow[r, "\epsilon_{\mathbf{1}}"'] & \mathbf{1}
\end{tikzcd}
\tag{5.3}$$

$$\begin{tikzcd}
FA \times FB \arrow[r, "\delta_A \times \delta_B"] \arrow[dd, "m_{A,B}"'] & F^2A \times F^2B \arrow[d, "m_{FA,FB}"] \\
 & F(FA \times FB) \arrow[d, "Fm_{A,B}"] \\
F(A \times B) \arrow[r, "\delta_{A \times B}"'] & F^2(A \times B)
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{1} \arrow[dr, "m_0"] \arrow[dd, "m_0"'] & & \\
 & F\mathbf{1} \arrow[dr, "Fm_0"] & \\
F\mathbf{1} \arrow[rr, "\delta_{\mathbf{1}}"'] & & F^2\mathbf{1}
\end{tikzcd}
\tag{5.4}$$

However, since the functors that we use are product-preserving, or strong monoidal, it follows automatically by Theorem 5.9 that

Corollary 5.14. *If* (F, ϵ, δ) *is a comonad whose functor* F *is product-preserving, then it is a monoidal comonad.*

We shall not hence explicitly worry about monoidality, neither in this section nor in later ones, and we will use the above equations without further ado.

Definition 5.15. A *Bierman-de Paiva category* $(\mathcal{C}, \times, \mathbf{1}, F, \epsilon, \delta)$ consists of a Kripke category $(\mathcal{C}, \times, \mathbf{1}, F)$ whose functor $F : \mathcal{C} \longrightarrow \mathcal{C}$ is part of a comonad (F, ϵ, δ).

Bierman-de Paiva categories (abbrv. BdP categories) are the minimal setting in which both the Four and Veridicality rules can be modelled. The Four rule is modelled by something already well-known in category theory, namely the *co-Kleisli lifting*:

$$(-)^* : \mathcal{C}\left(\prod_{i=1}^{n} FA_i, B\right) \to \mathcal{C}\left(\prod_{i=1}^{n} FA_i, FB\right)$$

which is defined as follows:

$$\frac{f : \prod_{i=1}^{n} FA_i \to B}{f^* \stackrel{\text{def}}{=} \prod_{i=1}^{n} FA_i \xrightarrow{\prod_{i=1}^{n} \delta_{A_i}} \prod_{i=1}^{n} F^2 A_i \xrightarrow{m^{(n)}} F\left(\prod_{i=1}^{n} FA_i\right) \xrightarrow{Ff} FB}$$

This operation interacts in a useful manner with the transformations δ and ϵ.

Proposition 5.16.
(1) *Let* $f : \prod_{i=1}^{n} FA_i \to B$. *Then* $\delta_B \circ f^* = (f^*)^*$.
(2) *Let* $f : \prod_{i=1}^{n} FA_i \to B$. *Then* $\epsilon_B \circ f^* = f$.

The co-Kleisli extension also satisfies a handful of very useful naturality/distribution laws.

Proposition 5.17.
(1) $id_{FA}^* = \delta_{FA}$
(2) $\epsilon_A^* = id_{FA}$
(3) *Let* $f : \prod_{i=1}^{n} B_i \to C$ *and* $g_i : \prod_{j=1}^{k} FA_j \to B_i$ *for* $i = 1, \dots, n$. *Then*

$$(f \circ \langle \overrightarrow{g_i} \rangle)^* = f^\bullet \circ \left\langle \overrightarrow{g_i^*} \right\rangle$$

(4) *For* $k_i : \prod_{j=1}^{m} FA_j \to B_j$ *and* $l : \prod_{i=1}^{n} FB_i \to C$,

$$\left(l \circ \left\langle \overrightarrow{k_i^*} \right\rangle\right)^* = l^* \circ \left\langle \overrightarrow{k_i^*} \right\rangle$$

(5) *For* $f : \prod_{i=1}^{n} FA_i \to B$ *and* $\langle \vec{\pi_j} \rangle : \prod_{i=1}^{n} FA_i \to \prod_{j \in J} FA_j$ *for* J *a list with elements from* $\{1, \dots, n\}$,

$$(f \circ \langle \vec{\pi_j} \rangle)^* = f^* \circ \langle \vec{\pi_j} \rangle$$

Proof. (1) and (2) are standard comonad equations. (3) is a straightforward calculation, similar to Proposition 5.11(1). (4) follows from (3), Proposition 5.16(1), and $f^* \stackrel{\text{def}}{=} f^\bullet \circ \prod \delta$. (5) is a corollary of (3), once we notice that $\pi_i^* = \delta_{A_i} \circ \pi_i$, and use $f^* \stackrel{\text{def}}{=} f^\bullet \circ \prod \delta$. □

Product-preserving comonads are also often encountered in the category-theoretic literature, so we refrain from sketching any examples of Bierman-de Paiva categories.

Idempotence. It is interesting to separately consider those BdP categories for which the comonad (F, ϵ, δ) is *idempotent*, i.e. those for which $\delta : F \Rightarrow F^2$ is an *isomorphism.* There are many equivalent ways to define idempotence: see [Bor94, §4.3.2]. One of them is that $\delta_{FA} \circ \epsilon_{FA} = id_{F^2A}$ for each object A. Here, we will use the equation $F\epsilon_A = \epsilon_{FA}$ for each object A. Restated in our notation, it becomes

$$\epsilon_A^\bullet = \epsilon_{FA} : F^2A \to FA$$

Proposition 5.18. *If (F, ϵ, δ) is idempotent, then for $f : \prod_{i=1}^n FA_i \to FB$ we have*

$$(\epsilon_B \circ f)^* = f$$

Proof. By Proposition 5.17(3) and 5.16(2), $(\epsilon \circ f)^* = \epsilon^\bullet \circ f^* = \epsilon \circ f^* = f$ □

This situation creates an additional sort of naturality for $(-)^*$, essentially by making Proposition 5.17(4) universally applicable.

Proposition 5.19. *For any $f : \prod_{j=1}^m FB_j \to FC$, and any $k_j : \prod_{i=1}^n FA_i \to FB_j$, we have*

$$\left(f \circ \left\langle \overrightarrow{k_i} \right\rangle\right)^* = f^* \circ \left\langle \overrightarrow{k_i} \right\rangle$$

Proof. Using Propositions 5.18 and 5.17(4),

$$\left(f \circ \left\langle \overrightarrow{k_j} \right\rangle\right)^* = \left(f \circ \left\langle \overrightarrow{(\epsilon_{B_j} \circ k_j)^*} \right\rangle\right)^* = f^* \circ \left\langle \overrightarrow{(\epsilon_{B_j} \circ k_j)^*} \right\rangle = f^* \circ \left\langle \overrightarrow{k_j} \right\rangle$$ □

Thus, we have the following characterisation.

Theorem 5.20 (Idempotence)**.** *(F, ϵ, δ) is idempotent if and only if the map*

$$(-)^* : \mathcal{C}\left(\prod_{i=1}^n FA_i, B\right) \to \mathcal{C}\left(\prod_{i=1}^n FA_i, FB\right)$$

is an isomorphism, natural with respect to precomposition of any $k : \prod_{j=1}^n FD_j \to \prod_{i=1}^n FA_i$.

Proof. For the backwards direction, the inverse of $(-)^*$ is $\epsilon_B \circ -$. It is a left and right inverse by Propositions 5.16(2) and 5.18. Naturality follows by Proposition 5.19 once we write $k = \left\langle \overrightarrow{\pi_i \circ k} \right\rangle$. For the forwards direction, we calculate

$$\delta_{FA} \circ \epsilon_{FA} = id_{FA}^* \circ \epsilon_{FA} = (id_{FA} \circ \epsilon_{FA})^* = \epsilon_{FA}^* = id_{F^2A}$$

by Prop. 5.17(1), naturality for $\epsilon_{FA} : F^2A \to FA$, and Prop. 5.17(2). □

Comparison with Bierman & de Paiva. Those familiar with previous literature on the categorical semantics of $\mathsf{S4}$ modalities will no doubt ask about the relationship of BdP categories with the models discussed by Bierman and de Paiva [BdP00, §7]. At their core, the models of the system with delayed substitutions that we discussed in §2.1 also consist of a monoidal comonad. However, the natural transformation $m : F(-) \times F(-) \Rightarrow F(- \times -)$ is not required to be invertible, and so F is only lax monoidal. As a result, one must also explicitly require the coherence equations (5.3) and (5.4), which—as we mentioned above—hold automatically whenever F preserves products.

In the penultimate section of their paper, Bierman and de Paiva [BdP00, §10] briefly discuss a slightly stronger notion of model, which they attribute to Schalk. Therein the morphisms $\langle id_{FA}, id_{FA} \rangle : FA \to FA \times FA$ and $!_{FA} : FA \to \mathbf{1}$ are homomorphisms for the coalgebras $\delta_A : FA \to F^2A$ and $m \circ (\delta_A \times \delta_A) : FA \times FA \to F^2A \times F^2A \to F(FA \times FA)$. These equations allow one to prove the soundness of two additional commuting conversions, which embody certain structural rules for the delayed substitutions: they allow one to *weaken* by 'garbage collecting' a delayed substitution for a variable that does not occur, and to *contract* two identical delayed substitutions. Indeed, these requirements also reappear in models of linear logic known as *linear categories* [Mel09, §7.4].

Curiously, our notion of model is even stronger: it is easy to calculate (using monoidality and naturality of δ, product preservation, and the invertibility of Fm) that the aforementioned morphisms are automatically coalgebra morphisms in the product-preserving setting. Indeed, two commuting conversions similar to the ones mentioned above, here called ($\mathsf{commweak}$) and ($\mathsf{commcontr}$), will be necessary to prove completeness.

5.4.3. *Kripke-4 categories.* Kripke-4 categories model $\mathsf{K4}$. They are essentially 'half a comonad,' and only come with a comultiplication δ. We still require that one of the comonadic equations, viz. the one that only refers to δ, holds.

Definition 5.21. A *Kripke-4* category $(\mathcal{C}, \times, \mathbf{1}, F, \delta)$ is a Kripke category $(\mathcal{C}, \times, \mathbf{1}, F)$ along with a natural transformation $\delta : F \Rightarrow F^2$ such that the following diagram commutes:

$$\begin{tikzcd} FA \arrow[r, "\delta_A"] \arrow[d, "\delta_A"'] & F^2A \arrow[d, "\delta_{FA}"] \\ F^2A \arrow[r, "F\delta_A"'] & F^3A \end{tikzcd}$$

We know by Theorem 5.9 that $\delta : F \Rightarrow F^2$ is a monoidal natural transformation.

We can model the general version of Four rule (1.5.2) in Kripke-4 categories, but in a way that is slightly more involved than the simple co-Kleisli lifting of Bierman–de Paiva categories. To see this, let $(\mathcal{C}, \times, \mathbf{1}, F, \delta)$ be a Kripke-4 category, and write

$$\prod_{i=1}^{n} A_i \times_l \prod_{j=1}^{m} B_j$$

to mean the left-associating product $A_1 \times \cdots \times A_n \times B_1 \times \cdots \times B_m$. Also, write

$$\langle \overrightarrow{f_i}, \overrightarrow{g_i}, \overrightarrow{h_j} \rangle$$

to mean the left-associating mediating morphism $\langle f_1, \dots, f_n, g_1, \dots, g_m, h_1, \dots, g_p \rangle$. With this notation we can now define a map of hom-sets

$$(-)^{\#} : \mathcal{C}\left(\prod_{i=1}^{n} FA_i \times_l \prod_{i=1}^{n} A_i, B\right) \to \mathcal{C}\left(\prod_{i=1}^{n} FA_i, FB\right)$$

as follows:

$$\frac{f : \prod_{i=1}^{n} FA_i \times_l \prod_{i=1}^{n} A_i \to B}{f^{\#} \stackrel{\text{def}}{=} \prod_{i=1}^{n} FA_i \xrightarrow{\langle \overrightarrow{\delta_{A_i}\pi_i}, \overrightarrow{\pi_i} \rangle} \prod_{i=1}^{n} F^2A_i \times_l \prod_{i=1}^{n} A_i \xrightarrow{m^{(2n)}} F\left(\prod_{i=1}^{n} FA_i \times_l \prod_{i=1}^{n} A_i\right) \xrightarrow{Ff} B}$$

Even though it might seem slightly contrived at first sight, we can show that the $(-)^{\#}$ operation satisfies some naturality equations similar to the ones encountered before.

Proposition 5.22.

(1) *Let* $f : \prod_{i=1}^n B_i \to C$ *and* $g_i : \prod_{j=1}^k FA_j \times_l \prod_{j=1}^k A_j \to B_i$ *for* $i = 1, \ldots, n$*. Then*

$$(f \circ \langle \overrightarrow{g_i} \rangle)^\# = f^\bullet \circ \left\langle \overrightarrow{g_i^\#} \right\rangle$$

(2) *Let* J *be a list with elements from* $\{1, \ldots, n\}$*. Then we have*

$$(f \circ \langle \overrightarrow{\pi_F j}, \overrightarrow{\pi_j} \rangle)^\# = f^\# \circ \langle \overrightarrow{\pi_j} \rangle$$

where $\langle \overrightarrow{\pi_F j}, \overrightarrow{\pi_j} \rangle : \prod_{i=1}^n FA_i \times_l \prod_{i=1}^n A_i \to \prod_{j \in J} FA_j \times_l \prod_{j \in J} A_j$ *is the projection that 'follows* J *in both contexts'* $\prod_{i=1}^n FA_i$ *and* $\prod_{i=1}^n A_i$*.*

(3) *For* $k : \prod_{i=1}^n FA_i \times_l \prod_{i=1}^n A_i \to B$ *and* $l : FB \to C$*, then*

$$(l \circ k^\#)^* = l^* \circ k^\#$$

Proof. Straightforward calculations, similar to Propositions 5.11 and 5.17. □

Proposition 5.23. *Let* $f : \prod_{i=1}^n FA_i \times_l \prod_{i=1}^n A_i \to B$*. Then*

$$\delta_B \circ f^\# = \left(f^\#\right)^*$$

When the morphism of type $\prod_{i=1}^n FA_i \times_l \prod_{i=1}^n A_i \to B$ does not depend on the A_i, then the operation $(-)^\#$ can be reduced to $(-)^*$.

Proposition 5.24. *Let* $f : \prod_{i=1}^n FA_i \to B$*. Then, writing* $\pi : \prod_{i=1}^n FA_i \times_l \prod_{i=1}^n A_i \to \prod_{i=1}^n FA_i$ *for the projection,*

$$(f \circ \pi)^\# = f^*$$

Proof. Straightforward calculation using Propositions 5.4 and 5.5. □

And, finally, we prove another crucial distribution property of $(-)^\#$. Namely, if we substitute 'the same thing' for both contexts, the hash distributes as such:

Proposition 5.25. *Let* $f : \prod_{i=1}^n FB_i \times \prod_{i=1}^n B_i \to B$*, and* $g_i : \prod_{i=1}^n FA_i \times \prod_{i=1}^n A_i \to B_i$*. Then, writing* $\pi : \prod_{i=1}^n FA_i \times \prod_{i=1}^n A_i \to \prod_{i=1}^n FA_i$ *for the projection,*

$$\left(f \circ \langle \overrightarrow{g_i^\# \circ \pi, \overrightarrow{g_i}} \rangle\right)^\# = f^\# \circ \left\langle \overrightarrow{g_i^\#} \right\rangle$$

Proof. Using Propositions 5.22(1), 5.24, and 5.23, we calculate

$$\left(f \circ \left\langle \overrightarrow{g_i^\# \circ \pi, \overrightarrow{g_i}} \right\rangle\right)^\# = f^\bullet \circ \left\langle \overrightarrow{\left(g_i^\# \circ \pi\right)^\#, \overrightarrow{g_i^\#}} \right\rangle = f^\bullet \circ \left\langle \overrightarrow{\left(g_i^\#\right)^*, \overrightarrow{g_i^\#}} \right\rangle = f^\bullet \circ \left\langle \overrightarrow{\delta \circ g_i^\#}, \overrightarrow{g_i^\#} \right\rangle$$

□

Example 5.26 (The topos of bifurcating trees, part 1)**.** The *topos of trees* [BMSS12] is the category $\mathbf{Psh}(\omega)$ of presheaves over $\omega \stackrel{\text{def}}{=} 1 < 2 < \ldots$. Concretely, each object X of $\mathbf{Psh}(\omega)$ is a diagram of sets and functions

$$X_0 \xleftarrow{r_0} X_1 \xleftarrow{r_1} X_2 \xleftarrow{r_2} \ldots$$

The idea is that X is a set that is computed over time: X_{i+1} contains the elements that can be emitted after computing for $i+1$ steps, and $r_i : X_{i+1} \to X_i$ *trims* such elements to what they were at i steps. The topos of trees is a synthetic model of step-indexed computation, as it provides a principled way of reasoning about infinite behaviours without ever constructing 'completed' objects: the latter only appear as global sections $x : \mathbf{1} \Rightarrow X$ of an object X, i.e. families $(x_i \in X_i)_i$ compatible under trimming.

A variation on this idea is the following. Let $(\mathbb{B}^*, \sqsubseteq)$ be the prefix order on words over booleans $\mathbb{B} \stackrel{\text{def}}{=} \{0, 1\}$. We can then construct the *topos of bifurcating trees*, namely the presheaf topos $\mathbf{Psh}(\mathbb{B}^*)$ over the prefix order. Its objects are diagrams

$$\begin{tikzcd}
 & & X_{00} \arrow[dl, "l_0"'] & \cdots \\
 & X_0 \arrow[dl, "l_\epsilon"'] & X_{01} \arrow[l, "r_0"] & \cdots \\
X_\epsilon & & & \\
 & X_1 \arrow[ul, "r_\epsilon"] & X_{10} \arrow[l, "l_1"'] & \cdots \\
 & & X_{11} \arrow[ul, "r_1"] & \cdots
\end{tikzcd}$$

Intuitively, each element of $x_w \in X_w$ might, in one time step, evolve in two ways: to an element $x_{w0} \in X_{w0}$, or to an element $x_{w1} \in X_{w1}$, such that $x_w = l_w(x_{w0}) = r_w(x_{w1})$. This encodes a certain degree of nondeterminism: a global section $x : \mathbf{1} \Rightarrow X$ now represents an infinite computation that may make a nondeterminsitic choice at every tick of the clock.

We may then perform the following construction: given a presheaf X over $\mathbb{B}^*$, we construct a presheaf $\Box X$ by letting

$$(\Box X)_w \stackrel{\text{def}}{=} \prod_{v \sqsubset w} X_v$$

where $\sqsubset$ is the *strict* order associated to the partial order $\sqsubseteq$. Thus, $(\Box X)_w$ consists of families $\{x_v\}_{v \sqsubset w}$ of an element for each word v that is a strict prefix of w. Whether those families are *matching*, i.e. whether for example $l_v(x_{v0}) = x_v$ whenever $v0 \sqsubset w$ is immaterial: the constructions sketched here work whether we read $\prod$ as a categorical limit or as a dependent product. This may prove to be yet another way in which we may obtain 'intensional' models of modal logic. The restriction maps $l_w : \prod_{v \sqsubset w1} X_v \to \prod_{v \sqsubset w} X_v$ and $r_w : \prod_{v \sqsubset w1} X_v \to \prod_{v \sqsubset w} X_v$ are defined by restricting the domain of $\prod$.

Furthermore, for each natural transformation $f : X \to Y$ we define $\Box f : \Box X \to \Box Y$ by

$$\begin{aligned}
(\Box f)_w : \prod_{v \sqsubset w} X_v &\to \prod_{v \sqsubset w} Y_v \\
p &\mapsto \lambda v.\ f_v(p(v))
\end{aligned}$$

$\mathbf{Psh}(\mathbb{B}^*)$ is a Kripke-4 category: we may define a $\delta : \Box \Rightarrow \Box^2$ at each X by

$$\begin{aligned}
\delta_{X,w} : \prod_{v \sqsubset w} X_v &\to \prod_{v \sqsubset w} \prod_{z \sqsubset v} Y_z \\
p &\mapsto \lambda v.\ \lambda z.\ p(z)
\end{aligned}$$

This is well-typed, as $\sqsubset$ is transitive and hence $z \sqsubset w$. This bears a strong likeness to Kripke semantics: the transitivity of the 'Kripke site' $(\mathbb{B}^*, \sqsubseteq)$ leads to a proof-relevant witness of the $\mathsf{4}$ axiom! In contrast, the 'non-reflexive' flavour of $\Box$ means that there is no way natural $\epsilon_X : \Box X \to X$: for each $w \in \mathbb{B}^*$, the component $\epsilon_{X,w}$ would have type $\prod_{v \sqsubset w} X_v \to X_w$, and there in general no way to produce an element of X_w given an element for each prefix of w.

5.4.4. *Kripke-T categories.* The following structure will be used to interpret axiom T.

Definition 5.27. A *Kripke-T* category $(\mathcal{C}, \times, \mathbf{1}, F, \epsilon)$ consists of a Kripke category $(\mathcal{C}, \times, \mathbf{1}, F)$ along with a natural transformation

$$\epsilon : F \Rightarrow \mathsf{Id}$$

Using Theorem 5.9 again we see $\epsilon : F \Rightarrow \mathsf{Id}$ is a monoidal natural transformation. Modelling veridicality rule (§1.5.2) amounts to precomposition with the product of a bunch of components of $\epsilon : F \Rightarrow \mathsf{Id}$. This operation interacts nicely with Scott's rule:

Proposition 5.28. *Let* $f : \prod_{i=1}^n A_i \to B$*. Then* $\epsilon_B \circ f^\bullet = f \circ \prod_{i=1}^n \epsilon_{A_i}$.

5.4.5. *Gödel-Löb categories.* We are looking for a setting where Löb's rule can be modelled. This is not as natural as the previous ones: it is closely modelled after the syntax of our calculus, and its definition will involve all of the operations $(-)^\bullet$, $(-)^\#$, and $(-)^*$.

We begin by defining the following central notion.

Definition 5.29 (Modal Fixed Point)**.** Let $(\mathcal{C}, \times, \mathbf{1}, F, \delta)$ be a Kripke-4 category.

(1) A *modal fixed point* of $f : \prod_{i=1}^n FB_i \times \prod_{i=1}^n B_i \times FA \to A$ is an arrow

$$f^\dagger : \prod_{i=1}^n FB_i \to FA$$

such that the following diagram commutes:

$$\begin{tikzcd}
\prod_{i=1}^n FB_i \arrow[r, "\langle id^\#, (f^\dagger)^* \rangle"] \arrow[d, "f^\dagger"'] & F\left(\prod_{i=1}^n FB_i \times \prod_{i=1}^n B_i\right) \times F^2 A \arrow[dl, "f^\bullet"] \\
FA &
\end{tikzcd}$$

(2) An object $A \in \mathcal{C}$ *has modal fixed points* just if for any $B_i \in \mathcal{C}$ there is a homset map

$$(-)^\dagger_{\vec{B}} : \mathcal{C}\left(\prod_{i=1}^n FB_i \times \prod_{i=1}^n B_i \times FA, A\right) \to \mathcal{C}\left(\prod_{i=1}^n FB_i, FA\right)$$

such that $f^\dagger_{\vec{B}}$ is a modal fixed point of each $f : \prod_{i=1}^n FB_i \times \prod_{i=1}^n B_i \times FA \to A$.

We will often write $f^\dagger$ for the modal fixed point of f, dropping the subscript entirely. This is an *external* specification of modal fixed points, in the sense that they are given as a map on the homsets of the Kripke-4 category. We might instead consider an *internal* specification, i.e. through an appropriate notion of a *modal fixed point combinator*. This will—unsurprisingly—be an arrow $F(A^{FA}) \to FA$, which is the type of the Gödel-Löb axiom $\Box(\Box A \to A) \to \Box A$. This kind of combinator comes in two varieties.

Definition 5.30. Let $(\mathcal{C}, \times, \mathbf{1}, F, \delta)$ be a Kripke-4 category.

(1) A *strong modal fixed point combinator at* $A \in \mathcal{C}$ is an arrow

$$Y_A : F(A^{FA}) \to FA$$

such that the following diagram commutes:

$$\begin{tikzcd} F(A^{FA}) \arrow[r, "\langle id, Y_A^* \rangle"] \arrow[d, "Y_A"'] & F(A^{FA}) \times F^2 A \arrow[dl, "\mathsf{ev}^\bullet"] \\ FA & \end{tikzcd}$$

(2) A *weak modal fixed point combinator at $A \in \mathcal{C}$* is an arrow

$$Y_A : F(A^{FA}) \to FA$$

such that for each B and $f : \prod_{i=1}^n FB_i \times \prod_{i=1}^n B_i \times FA \to A$, the composite

$$\prod_{i=1}^n FB_i \xrightarrow{(\lambda(f))^\#} F(A^{FA}) \xrightarrow{Y_A} FA$$

is a modal fixed point of f.

We can prove that having a modal fixed point combinator at A is equivalent to having modal fixed points at A. But to do so we will need a lemma concerning cartesian closure.

Lemma 5.31. *If $f : \prod_{i=1}^n FA_i \times \prod_{i=1}^n A_i \times FB \to B$ and $a : \prod_{i=1}^n FA_i \to F^2 A$, then*

$$\mathsf{ev}^\bullet \circ \langle (\lambda f)^\# , a \rangle = f^\bullet \circ \langle id^\#, a \rangle$$

Theorem 5.32. *Let $(\mathcal{C}, \times, \mathbf{1}, F, \delta)$ be a Kripke-4 category. The following are equivalent:*
(1) *There is a strong modal fixed point combinator at A.*
(2) *There is a weak modal fixed point combinator at A.*
(3) *The object $A \in \mathcal{C}$ has modal fixed points.*

Proof. To prove (1) $\Rightarrow$ (2), if we are given such a Y we calculate

$$\begin{aligned} & Y \circ \lambda f^\# \\ = \quad & \{ \text{definition of strong mfpc, naturality of product morphism} \} \\ & \mathsf{ev}^\bullet \circ \langle \lambda f^\#, Y^* \circ \lambda f^\# \rangle \\ = \quad & \{ \text{Proposition 5.22(3)} \} \\ & \mathsf{ev}^\bullet \circ \langle \lambda f^\#, (Y \circ \lambda f^\#)^* \rangle \\ = \quad & \{ \text{Proposition 5.31} \} \\ & f^\bullet \circ \langle id^\#, (Y \circ \lambda f^\#)^* \rangle \end{aligned}$$

so Y yields modal fixed points. (2) $\Rightarrow$ (3) is trivial, so it remains to show (3) $\Rightarrow$ (1). Let

$$g \stackrel{\text{def}}{=} F(A^{FA}) \times A^{FA} \times FA \xrightarrow{\langle \pi_2, \pi_3 \rangle} A^{FA} \times FA \xrightarrow{\mathsf{ev}} A$$

We show that $g^\dagger : F(A^{FA}) \to FA$ is a strong modal fixed point combinator at A. Indeed, it is not very hard to calculate that

$$g^\dagger = \mathsf{ev}^\bullet \circ \langle id, \left(g^\dagger\right)^* \rangle \qquad \square$$

We formulate the following naturality property of modal fixed points, which is partly reminiscent of the ones of Simpson and Plotkin [SP00], but also resembles Proposition 5.25.

Proposition 5.33. *If we define $(-)^\dagger_B$ by a weak modal fixed point combinator, then the resulting modal fixed points are* natural, *in the sense that for any $f : \prod_{i=1}^n FA_i \times \prod_{i=1}^n A_i \times A \to A$ and any $g_i : \prod_{j=1}^m FB_j \times \prod_{j=1}^m B_j \to B_i$, then, writing $\pi : \prod_{j=1}^m FB_i \times \prod_{j=1}^m A_j \to \prod_{j=1}^m FB_i$ for the projection,*

$$\left(f \circ \left(\left\langle \overrightarrow{g_i^\# \circ \pi}, \overrightarrow{g_i^\#} \right\rangle \times id_A\right)\right)^\dagger = f^\dagger \circ \left\langle \overrightarrow{g_i^\#} \right\rangle$$

Proof. Recall that $f^\dagger \stackrel{\text{def}}{=} Y \circ (\lambda(f))^\#$. The LHS is equal to

$$Y \circ \left(\lambda \left(f \circ \left(\left\langle \overrightarrow{g_i^\# \circ \pi}, \overrightarrow{g_i^\#} \right\rangle \times id_A\right)\right)\right)^\#$$

which, using naturality of $\lambda(-)$ and Proposition 5.25, is equal to $Y \circ (\lambda(f))^\# \circ \left\langle \overrightarrow{g_i^\#} \right\rangle$. □

We are finally in a position to define the notion of model used for GL.

Definition 5.34. A *Gödel-Löb category* $\left(\mathcal{C}, \times, \mathbf{1}, F, \delta, (-)^\dagger\right)$ is a Kripke-4 category $(\mathcal{C}, \times, \mathbf{1}, F, \delta)$ that has modal fixed points at all objects A, given by maps

$$(-)^\dagger_{\vec{B},A} : \mathcal{C}\left(\prod_{i=1}^n FB_i \times \prod_{i=1}^n B_i \times FA, A\right) \to \mathcal{C}\left(\prod_{i=1}^n FB_i, FA\right)$$

which, moreover, are *natural*, in the sense that for any $f : \prod_{i=1}^n FB_i \times \prod_{i=1}^n B_i \times A \to A$ and $g_i : \prod_{j=1}^m FC_j \times \prod_{j=1}^m C_j \to B_i$,

$$\left(f \circ \left(\left\langle \overrightarrow{g_i^\# \circ \pi}, \overrightarrow{g_i^\#} \right\rangle \times id_A\right)\right)^\dagger_{\vec{C},A} = f^\dagger_{\vec{B},A} \circ \left\langle \overrightarrow{g_i^\#} \right\rangle$$

Combining the preceding theorem and proposition assures us that, we see that it does not matter how modal fixed points are given, as we can always turn them into a standard Gödel-Löb category.

We also show that, whenever the modal fixed point has no 'diagonal' occurences, it deteriorates to the $(-)^\#$ operation.

Proposition 5.35. *If $f : \prod_{i=1}^n FA_i \times \prod_{i=1}^n A_i \to B$ and $\pi : \prod_{i=1}^n FA_i \times \prod_{i=1}^n A_i \times A \to \prod_{i=1}^n FA_i \times \prod_{i=1}^n A_i$ is the obvious projection, then*

$$(f \circ \pi)^\dagger = f^\#$$

Proof. Writing $g \stackrel{\text{def}}{=} (f \circ \pi)^\dagger$, we have

$$g = (f \circ \pi)^\bullet \circ \left\langle id^\#, g^* \right\rangle = f^\bullet \circ \pi \circ \left\langle id^\#, g^* \right\rangle = f^\bullet \circ id^\# = (f \circ id)^\# = f^\#$$

by the definition of modal fixed point, and Prop. 5.11(2), 5.22(1). □

Example 5.36 (The topos of bifurcating trees, part 2)**.** Recall Example 5.26. The topos of trees is a model of *guarded recursion*. Define the functor $\blacktriangleright : \mathbf{Psh}(\omega) \to \mathbf{Psh}(\omega)$ by having it 'delay' a computation by one time step, i.e. by mapping $X_1 \xleftarrow{r_1} X_2 \xleftarrow{r_2} \dots$ to $\mathbf{1} \xleftarrow{!} X_1 \xleftarrow{r_1} \dots$. This admits a natural transformation $\mathsf{next}_X : X \to \blacktriangleright X$ which 'delays' a computation by trimming $x_{n+1} \in X_{n+1}$ to $r_n(x_{n+1}) \in X_n = (\blacktriangleright X)_{n+1}$. We may perform guarded recursion: given $f : \blacktriangleright X \Rightarrow X$, we define a global section $x : \mathbf{1} \Rightarrow X$ by

$$x_1 \stackrel{\text{def}}{=} f_1(*) : X_1 \qquad\qquad x_{n+1} \stackrel{\text{def}}{=} f_{n+1}(x_n) : X_{n+1}$$

Essentially $f : \blacktriangleright X \to X$ provides both a 'seed' value $f_1(*)$ as well as a 'coinductive step function' $f_{n+1} : X_n \to X_{n+1}$ at each tick of the clock. We thus have *guarded fixed points*, in that $x = f \circ \mathsf{next}_X \circ x$. This can be done internally, so it defines a map $\mathsf{fix}_X : (\blacktriangleright A \to A) \to A$ corresponding to the strong Gödel-Löb logic axiom of the logic SL.

A similar construction can be carried out for the topos of bifurcating trees. We can define a natural transformation $\mathsf{erstwhile}_X : X \to \Box X$ by

$$\begin{aligned} \mathsf{erstwhile}_{X,w} : X_w &\to \prod_{v \sqsubset w} X_v \\ x_w &\mapsto \lambda v.\ x_w|_v \end{aligned}$$

where $(-)|_v$ is the obvious presheaf action $X_w \to X_v$ that trims an element at 'stage' w to an element at 'stage' v for any $v \sqsubset w$. A natural transformation $f : \Box X \Rightarrow X$ is a collection of maps $f_w : \prod_{v \sqsubset w} X_v \to X_w$. In short, f witnesses a proof-relevant *strong induction hypothesis*: when given witnesses of X at each stage that strictly precedes w, it returns a witness of X at stage w. We can use this witness to define $x : \mathbf{1} \Rightarrow X$ by

$$x_w \stackrel{\text{def}}{=} f_w(\lambda v \sqsubset w.\ x_v)$$

This definition is admissible precisely because f_w only depends on x_v for a *strict* prefix v, and because the prefix order is *well-founded* [HJ99, §14.1]. This closely mirrors the situation from the classical Kripke semantics of GL, which is sound and complete for transitive frames for which the converse of the accessibility relation is *well-founded* [Boo94, §4]. We have $x = f \circ \mathsf{erstwhile}_X \circ x$.

By slightly generalising this construction we can show that it furnishes something stronger than the Gödel-Löb axiom: like the topos of trees, it is a proof-relevant model of SL, i.e. a *guarded fixpoint category* in the sense of Milius and Litak [ML13]:

Definition 5.37 (Guarded Fixpoint Category)**.** A *guarded fixpoint category* $(\mathcal{C}, \blacktriangleright, p, (-)^{\pm})$ consists of a category $\mathcal{C}$ with finite products, an endofunctor $\blacktriangleright : \mathcal{C} \to \mathcal{C}$, a natural transformation $p : \mathsf{Id} \Rightarrow \blacktriangleright$, and a map of homsets

$$(-)^{\pm}_{B,A} : \mathcal{C}(B \times \blacktriangleright A, A) \to \mathcal{C}(B, A)$$

such that for each $f : B \times \blacktriangleright A \to A$, the morphism $f^{\pm}_{B,A} : B \to A$ is a *guarded fixpoint of* f, i.e. a morphism for which the following diagram commutes:

$$\begin{tikzcd} B \arrow[r, "f^{\pm}"] \arrow[d, "\langle id, f^{\pm} \rangle"'] & A \\ B \times A \arrow[r, "id \times p_A"'] & B \times \blacktriangleright A \arrow[u, "f"'] \end{tikzcd}$$

Indeed, Milius and Litak mention that any presheaf topos over *any* well-founded order forms a guarded fixpoint category [ML13, Example 2.4(5)]. In fact, any such category is also a Gödel-Löb category.

Theorem 5.38. *Let $(\mathcal{C}, F, p, (-)^{\pm})$ be a guarded fixpoint category, where F preserves finite products. Then $(\mathcal{C}, \times, \mathbf{1}, F, Fp)$ is a Kripke-4 category, which may be equipped with the structure of a Gödel-Löb strategy by defining the modal fixed point of each $f : \prod_{i=1}^{n} FB_i \times \prod_{i=1}^{n} B_i \times FA \to A$ to be*

$$f^{\dagger} \stackrel{\textit{def}}{=} \left(f^{\pm}\right)^{\#} : \prod_{i=1}^{n} FB_i \to FA$$

Proof. A quick calculation shows that $Fp : F \Rightarrow F^2$ satisfies Definition 5.21. That $f^{\dagger}$ is a modal fixed point is a straightforward calculation using Props. 5.22(1) and 5.23. □

We still do not know whether there are any interesting Gödel-Löb categories that do not arise from guarded fixpoint categories via the above.

6. Categorical semantics

In this section we use the modal category theory developed in §5 to formulate a categorical semantics for our dual-context calculi. This completes the circle in terms of the Curry-Howard-Lambek correspondence by establishing the following associations:

CK	$\longleftrightarrow$	DK	$\longleftrightarrow$	Kripke categories
$\mathsf{CK4}$	$\longleftrightarrow$	$\mathsf{DK4}$	$\longleftrightarrow$	Kripke-4 categories
CGL	$\longleftrightarrow$	DGL	$\longleftrightarrow$	Gödel-Löb categories
CT	$\longleftrightarrow$	DT	$\longleftrightarrow$	Kripke-T categories
$\mathsf{CS4}$	$\longleftrightarrow$	$\mathsf{DS4}$	$\longleftrightarrow$	Bierman-de Paiva categories

where the first bi-implication refers to provability, and the second to soundness and completeness of the dual-context calculus with respect to the categorical model on the right.

We begin by endowing our calculi with an equational theory. We then propose a categorical interpretation, and show that it is sound. Finally, we discuss completeness.

6.1. Equational theory.

Our equational theory of modal proofs should at the very least contain the reductions used in §4. It should also come with η rules, which we did not include in §4 due to their usual problematic behaviour under reduction.

A fragment of the equational theory may be found in Figure 6. To generate the theory for, say, DGL, we take the first three rules (β/η for function types, η for modal types), as well as the appropriate β rule for each system—in this case, $(\Box\beta_{\mathsf{GL}})$. To these we must not forget to include (a) rules that ensure that equality is an equivalence relation, and (b) the usual congruence rules for all term formers. The congruence rules for $\mathsf{box}\ (-)$ must be typed with care. For example, the congruence rule for $\mathsf{DK4}$ should be

$$\frac{\Delta\,;\Delta^{\perp} \vdash M^{\perp} = N^{\perp} : A}{\Delta\,;\Gamma \vdash \mathsf{box}\ M = \mathsf{box}\ N : \Box A}$$

We need not include substitution rules:

Theorem 6.1. *Structural rules of weakening, exchange and contraction for contexts are admissible in the equational theory. Furthermore, the following rules are derivable:*

(1) *Substitution:*

$$\frac{\Delta\,;\Gamma, x : A \vdash M = N : C \qquad \Delta\,;\Gamma \vdash P = Q : A}{\Delta\,;\Gamma \vdash M[P/x] = N[Q/x] : C}$$

$$\frac{\Delta \,;\Gamma, x : A \vdash M : B \qquad \Delta \,;\Gamma \vdash N : A}{\Delta \,;\Gamma \vdash (\lambda x : A.M)\, N = M[N/x] : B}\,(\to\beta) \qquad \frac{\Delta \,;\Gamma \vdash M : A \to B \qquad x \notin \mathrm{FV}\,(M)}{\Delta \,;\Gamma \vdash M = \lambda x : A.Mx : A \to B}\,(\to\eta)$$

$$\frac{\Delta \,;\Gamma \vdash M : \Box A}{\Delta \,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in\ box}\ u = M : \Box A}\,(\Box\eta)$$

$$\frac{\cdot \,;\Delta \vdash M : A \qquad \Delta, u : A \,;\Gamma \vdash N : C}{\Delta \,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ M\ \mathsf{in}\ N = N[M/x] : C}\,(\Box\beta_{\mathsf{K}})$$

$$\frac{\Delta \,;\Delta^{\perp} \vdash M^{\perp} : A \qquad \Delta, u : A \,;\Gamma \vdash N : C}{\Delta \,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ M\ \mathsf{in}\ N = N[M/x] : C}\,(\Box\beta_{\mathsf{K4}})$$

$$\frac{\Delta \,;\Delta^{\perp}, z^{\perp} : \Box A \vdash M^{\perp} : A \qquad \Delta, u : A \,;\Gamma \vdash N : C}{\Delta \,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow \mathsf{fix}\ z\ \mathsf{in\ box}\ M\ \mathsf{in}\ N = N\left[M\left[\mathsf{fix}\ z\ \mathsf{in\ box}\ M/z\right]/u\right] : C}\,(\Box\beta_{\mathsf{GL}})$$

$$\frac{\Delta \,;\cdot \vdash M : A \qquad \Delta, u : A \,;\Gamma \vdash N : C}{\Delta \,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ M\ \mathsf{in}\ N = N[M/x] : C}\,(\Box\beta_{\mathsf{S4}})$$

Figure 6: Equational theory

$$\frac{\Delta \,;\Gamma \vdash N : C \qquad \Delta \,;\Gamma \vdash M : \Box A \qquad u \notin \mathrm{FV}\,(N)}{\Delta \,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N = N : C}\,(\mathsf{commweak})$$

(commcontr):

$$\frac{\Delta \,;\Gamma \vdash M : \Box A \qquad \Delta, u : A, v : A \,;\Gamma \vdash N : C \qquad u, v \notin \mathrm{FV}\,(M)}{\Delta \,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in\ let\ box}\ v \Leftarrow M\ \mathsf{in}\ N = \mathsf{let\ box}\ w \Leftarrow M\ \mathsf{in}\ N[w, w/u, v] = N : C}$$

$$\frac{\Delta \,;\Gamma \vdash C[\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N] : C \qquad C[-]\text{ is non-modal, does not bind } u}{\Delta \,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ C[N] = C[\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N] : C}\,(\mathsf{commlet})$$

Figure 7: Commuting conversions

(2) *Modal Substitution: for example, in the case of* DK*:*

$$\frac{\Delta, u : A \,;\Gamma \vdash M = N : C \qquad \cdot\,;\Delta \vdash P = Q : \Box A}{\Delta \,;\Gamma \vdash M[P/u] = N[Q/u] : C}$$

Commuting Conversions. The most interesting rules are the unavoidable commuting conversions that we need if we want our categorical semantics to be *complete*. To state these we will need the notion of *term contexts*, i.e. terms with a single *hole*.

Definition 6.2 (Term Contexts)**.**

(1) Term contexts are generated by the grammar

$$\begin{aligned} C[-] \ ::= \ & [-] \mid \lambda x : A.\ C[-] \mid C[-]\ M \mid M\ C[-] \mid \langle C[-], M\rangle \mid \langle M, C[-]\rangle \mid \pi_i(C[-]) \\ & \mid \mathsf{box}\ C[-] \mid \mathsf{let\ box}\ u \Leftarrow C[-]\ \mathsf{in}\ M \mid \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ C[-] \end{aligned}$$

(2) A term context $C[-]$ is *non-modal* just if it is generated without the clause $\mathsf{box}\ C[-]$.
(3) $C[-]$ does not bind u just if its generation uses neither $\mathsf{let\ box}\ u \Leftarrow C[-]\ \mathsf{in}\ M$ nor $\lambda u : A.\, C[-]$.

We write $C[M]$ for the term that results from capture-insensitive substitution of the term M for the hole $[-]$ of the term context $C[-]$.

Our systems share the same set of commuting conversions, which may be found in Figure 7. The rule ($\mathsf{commweak}$) is a 'weakening,' or 'garbage collection' rule that disposes of a delayed substitution that binds a non-occurring variable. This rule has never been considered in the study of dual-context systems, for DILL [Bar96] was a linear system, and Davies and Pfenning [PD01] did not study reduction, equality, or categorical semantics. However, a similar rule was proposed by Goubault-Larrecq [GL96] in his study of Bierman and de Paiva's calculus for $\mathsf{S4}$. This rule was later included in [BdP00].

Similarly, ($\mathsf{commcontr}$) is a 'contraction' rule. This is also unfamiliar in dual-context calculi—essentially for the same reasons as ($\mathsf{commweak}$)—but is also well-known in Bierman–de Paiva style calculi as a 'garbage collection' rule: see [GL96], [BdP00] and [Kak07].

'Exchange' is treated as part of the much more general rule ($\mathsf{commlet}$), which makes 'let' constructs commute with all term formers except $\mathsf{box}\ (-)$. The equality is

$$C[\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N] = \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ C[N]$$

for any context C that does not bind u, and whose hole $[-]$ is not included within a $\mathsf{box}\ (-)$. Read in one direction, ($\mathsf{commlet}$) allows one to 'pull' a delayed substitution to an outermost position, as long as nothing extra is bound in the process. In the other direction, it allows one to 'push' a delayed substitution as deeply as one can without creating any free occurrences. A variant of this rule for DILL was considered by [Bar96], and is also mentioned by [Kak07].

The η rule. As is usual with *positive* type formers, of which $\Box$ is an example, there are two ways to express the η rule. The first is the straightforward way, viz. that introduction is post-inverse to elimination: for any term $\Delta\,;\Gamma \vdash M : \Box A$,

$$\Delta\,;\Gamma \vdash M = \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in\ box}\ u : \Box A$$

The second version of that is an *extended η-rule*, which allows us to η-expand a term of modal type, no matter where it is found in a well-typed term:

$$\frac{\Delta\,;\Gamma \vdash M : \Box A \qquad \Delta\,;\Gamma, x : \Box A \vdash N : B}{\Delta\,;\Gamma \vdash N[M/x] = \mathsf{let\ box}\ v \Leftarrow M\ \mathsf{in}\ N[\mathsf{box}\ v/x] : B}$$

In fact,

Theorem 6.3. *The η rule and the extended η-rule for the modal type are equivalent in the presence of commuting conversions.*

Proof. Certainly the η rule is a special case of the extended η-rule. In the opposite direction, we proceed by induction on the derivation of the term N. Most cases are simple. For products we have that

$$
\begin{aligned}
& \langle N_1, N_2\rangle[M/x] \\
\equiv\ & \{\ \text{substitution}\ \} \\
& \langle N_1[M/x], N_2[M/x]\rangle \\
=\ & \{\ \text{IH, twice}\ \} \\
& \langle \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N_1[\mathsf{box}\ u/x], \mathsf{let\ box}\ v \Leftarrow M\ \mathsf{in}\ N_2[\mathsf{box}\ v/x]\rangle \\
=\ & \{\ (\mathsf{commlet}),\ \text{twice}\ \} \\
& \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ \mathsf{let\ box}\ v \Leftarrow M\ \mathsf{in}\ \langle N_1[\mathsf{box}\ u/x], N_2[\mathsf{box}\ v/x]\rangle \\
=\ & \{\ (\mathsf{commcontr})\ \} \\
& \mathsf{let\ box}\ w \Leftarrow M\ \mathsf{in}\ \langle N_1[\mathsf{box}\ w/x], N_2[\mathsf{box}\ w/x]\rangle \\
\equiv\ & \{\ \text{substitution}\ \} \\
& \mathsf{let\ box}\ w \Leftarrow M\ \mathsf{in}\ N[\mathsf{box}\ w/x]
\end{aligned}
$$

A similar 'collapsing step' is also needed in the case of let. The case for $\mathsf{box}\ M$ is simple, as in all of our type theories x does not occur in M; the result hence follows by ($\mathsf{commweak}$). □

Idempotence in DS4. The ($\mathsf{commlet}$) rule avoided instances of commutation between a let and a box. If such commutations were allowed we would have for example the following equality in DS4:

$$\Delta\,;\Gamma \vdash \mathsf{box}\ (\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N) = \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in\ box}\ N : C$$

for $\Delta\,;\cdot \vdash M : \Box A$ and $\Delta, u : A\,;\cdot \vdash N : C$. We will later show that these rules are sound for the categorical semantics of DS4 if and only if the comonad used to interpret $\Box$ is *idempotent*.

There are three equivalent ways to present idempotence in DS4. The first two roughly say that $\mathsf{box}\ (-)$ and let commute. The third is a strong form of the extended η-rule, which this time applies to modal variables. Variants of this rule are sometimes known as *crisp induction* [Shu18, §5].

Theorem 6.4. *The following rules are equivalent*

(1)

$$\frac{\Delta\,;\cdot \vdash M : \Box A \quad \Delta, u : A\,;\cdot \vdash N : B}{\Delta\,;\Gamma \vdash \mathsf{box}\ (\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N) = \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in\ box}\ N : \Box B}$$

(2)

$$\frac{\Delta\,;\Gamma \vdash C[\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N] : B \qquad C[-]\ \textit{does not bind}\ u}{\Delta\,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ C[N] = C[\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N] : B}$$

(3)

$$\frac{\Delta\,;\cdot \vdash M : \Box A \quad \Delta, u : \Box A\,;\Gamma \vdash N : B}{\Delta\,;\Gamma \vdash N[M/u] = \mathsf{let\ box}\ v \Leftarrow M\ \mathsf{in}\ N[\mathsf{box}\ v/u] : B}$$

Proof. (1) is a special case of (2). To prove (2) from (1), we proceed by induction on C: use the commuting conversion ($\mathsf{commlet}$) for the non-modal cases, and then (1) for the modal case $C[-] \stackrel{\text{def}}{=} \mathsf{box}\ C'[-]$. If we have the premises of (1), we can show that, by (3),

$$(\mathsf{box}\ (\mathsf{let\ box}\ u \Leftarrow v\ \mathsf{in}\ N))\,[M/v] : \Box B$$

is equal to

$$\mathsf{let\ box}\ w \Leftarrow M\ \mathsf{in}\ \left(\mathsf{box}\ (\mathsf{let\ box}\ u \Leftarrow v\ \mathsf{in}\ N)\right)[\mathsf{box}\ w/v] : \Box B$$

The first expression simplifies to $\mathsf{box}\ (\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N)$, and the second to

$$\mathsf{let\ box}\ w \Leftarrow M\ \mathsf{in}\ \left(\mathsf{box}\ (\mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ v\ \mathsf{in}\ N)\right)$$

which, by one step of β-reduction and α-conversion is equal to $\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in\ box}\ N$. We can show (3) from (1) by induction on the derivation of N as before, but using (1) for the crucial case of $\mathsf{box}\ N'$. □

6.2. **Categorical interpretation.** We are now fully equipped to define the categorical semantics of our dual-context systems. For background on the categorical semantics of simply-typed λ-calculus in cartesian closed categories, we refer the reader to [LS88, Cro93, AT11].

We start by interpreting types and contexts. Given any Kripke category $(\mathcal{C}, \times, \mathbf{1}, F)$, and a map $\mathcal{I}(-)$ associating each base type p_i with an object $\mathcal{I}(p_i) \in \mathcal{C}$, we define an object $[\![A]\!] \in \mathcal{C}$ for every type A by induction:

$$[\![p_i]\!] \stackrel{\text{def}}{=} \mathcal{I}(p_i) \qquad\qquad [\![A \times B]\!] \stackrel{\text{def}}{=} [\![A]\!] \times [\![B]\!]$$

$$[\![A \to B]\!] \stackrel{\text{def}}{=} [\![B]\!]^{[\![A]\!]} \qquad\qquad [\![\Box A]\!] \stackrel{\text{def}}{=} F\,[\![A]\!]$$

Then, given a context $\Delta\,;\Gamma$ where $\Delta = u_1 : B_1, \dots u_n : B_n$ and $\Gamma = x_1 : A_1, \dots, x_m : A_m$, we let

$$[\![\Delta\,;\Gamma]\!] \stackrel{\text{def}}{=} FB_1 \times \cdots \times FB_n \times A_1 \times \cdots \times A_m$$

where the product is, as ever, left-associating. We then extend $[\![-]\!]$ to associate an arrow

$$[\![\Delta\,;\Gamma \vdash M : A]\!] : [\![\Delta\,;\Gamma]\!] \to [\![A]\!]$$

of the category $\mathcal{C}$ to each derivation $\Delta\,;\Gamma \vdash M : A$. The definition for rules common to all calculi are the same for all logics, but we use each of the maps defined in §5 to interpret the different introduction rules for the modality. To do that we need the corresponding structure we introduced in §5, e.g. for $\mathsf{K4}$ we need a Kripke-4 category, and so on.

The full definition is given in Figure 8. The morphism $\pi^{\Delta;\Gamma}_{\Delta} : [\![\Delta\,;\Gamma]\!] \to [\![\Delta\,;\cdot]\!]$ is the obvious projection. Moreover, the notation $\langle \overrightarrow{\pi_\Delta}, f, \overrightarrow{\pi_\Gamma} \rangle$ stands for

$$\langle \overrightarrow{\pi_\Delta}, f, \overrightarrow{\pi_\Gamma} \rangle \stackrel{\text{def}}{=} \langle \pi_1, \dots, \pi_n, f, \pi_{n+1}, \dots, \pi_{n+m} \rangle$$

6.3. **Soundness.** The main tools used in proving soundness are lemmas giving the categorical interpretation of various admissible rules, and a fundamental result relating substitution of terms to composition in the category. In the sequel we often use informal vector notation for contexts: for example, we write $\vec{u} : \vec{B}$ for the context $u_1 : B_1, \dots, u_m : B_m$. We also write $[\vec{N}/\vec{u}]$ for the simultaneous capture-avoiding substitution $[N_1/u_1, \dots, N_m/u_m]$.

First, we interpret weakening and exchange.

Lemma 6.5 (Semantics of Weakening)**.**

(1) *Let* $\Delta\,;\Gamma, x : C, \Gamma' \vdash M : A$ *with* $x \notin \mathrm{FV}(M)$. *Then*

$$\left[\!\!\left[\Delta\,;\Gamma, x : C, \Gamma' \vdash M : A\right]\!\!\right] = \left[\!\!\left[\Delta\,;\Gamma, \Gamma' \vdash M : A\right]\!\!\right] \circ \pi$$

where $\pi : [\![\Delta\,;\Gamma, x : C, \Gamma']\!] \to [\![\Delta\,;\Gamma, \Gamma']\!]$ *is the obvious projection.*

Definitions for all calculi

$$\llbracket \Delta\,;\Gamma, x{:}A, \Gamma' \vdash x : A \rrbracket \stackrel{\text{def}}{=} \pi : \llbracket \Delta\,;\Gamma, x{:}A, \Gamma' \rrbracket \longrightarrow \llbracket A \rrbracket$$
$$\llbracket \Delta\,;\Gamma \vdash \langle M, N\rangle : A \times B \rrbracket \stackrel{\text{def}}{=} \langle \llbracket \Delta\,;\Gamma \vdash M : A \rrbracket, \llbracket \Delta\,;\Gamma \vdash N : B \rrbracket \rangle$$
$$\llbracket \Delta\,;\Gamma \vdash \pi_i(M) : A_i \rrbracket \stackrel{\text{def}}{=} \pi_i \circ \llbracket \Delta\,;\Gamma \vdash M : A_1 \times A_2 \rrbracket$$
$$\llbracket \Delta\,;\Gamma \vdash \lambda x : A.\ M : A \to B \rrbracket \stackrel{\text{def}}{=} \lambda\left(\llbracket \Delta\,;\Gamma, x : A \vdash M : B \rrbracket\right)$$
$$\llbracket \Delta\,;\Gamma \vdash MN : B \rrbracket \stackrel{\text{def}}{=} \mathsf{ev} \circ \langle \llbracket \Delta\,;\Gamma \vdash M : A \to B \rrbracket, \llbracket \Delta\,;\Gamma \vdash N : A \rrbracket \rangle$$
$$\llbracket \Delta\,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N : C \rrbracket \stackrel{\text{def}}{=} \llbracket \Delta, u : A\,;\Gamma \vdash N : C \rrbracket \circ \langle \overrightarrow{\pi_\Delta}, \llbracket \Delta\,;\Gamma \vdash M : \Box A \rrbracket, \overrightarrow{\pi_\Gamma} \rangle$$

Definitions for various modalities

$$\llbracket \Delta, u : A, \Delta'\,;\Gamma \vdash u : A \rrbracket_{\mathcal{L}} \stackrel{\text{def}}{=} \epsilon_{\llbracket A \rrbracket} \circ \pi : \llbracket \Delta, u : A, \Delta'\,;\Gamma \rrbracket \to \llbracket A \rrbracket \qquad (\text{for } \mathcal{L} \in \{\mathsf{T}, \mathsf{S4}\})$$
$$\llbracket \Delta\,;\Gamma \vdash \mathsf{box}\ M : \Box A \rrbracket_{\mathcal{L}} \stackrel{\text{def}}{=} \llbracket \cdot\,;\Delta \vdash M : A \rrbracket^{\bullet} \circ \pi_{\Delta}^{\Delta;\Gamma} \qquad (\text{for } \mathcal{L} \in \{\mathsf{K}, \mathsf{T}\})$$
$$\llbracket \Delta\,;\Gamma \vdash \mathsf{box}\ M : \Box A \rrbracket_{\mathsf{K4}} \stackrel{\text{def}}{=} \left\llbracket \Delta\,;\Delta^{\perp} \vdash M^{\perp} : A \right\rrbracket^{\#} \circ \pi_{\Delta}^{\Delta;\Gamma}$$
$$\llbracket \Delta\,;\Gamma \vdash \mathsf{fix}\ z\ \mathsf{in\ box}\ M : \Box A \rrbracket_{\mathsf{GL}} \stackrel{\text{def}}{=} \left\llbracket \Delta\,;\Delta^{\perp}, z^{\perp} : \Box A \vdash M^{\perp} : A \right\rrbracket^{\dagger} \circ \pi_{\Delta}^{\Delta;\Gamma}$$
$$\llbracket \Delta\,;\Gamma \vdash \mathsf{box}\ M : \Box A \rrbracket_{\mathsf{S4}} \stackrel{\text{def}}{=} \llbracket \Delta\,;\cdot \vdash M : A \rrbracket^{*} \circ \pi_{\Delta}^{\Delta;\Gamma}$$

Figure 8: Categorical Semantics

(2) *Let* $\Delta, u : B, \Delta'\,;\Gamma \vdash M : A$ *with* $u \notin \mathrm{FV}(M)$. *Then*
$$\llbracket \Delta, u : B, \Delta'\,;\Gamma \vdash M : A \rrbracket = \llbracket \Delta, \Delta'\,;\Gamma \vdash M : A \rrbracket \circ \pi$$
where $\pi : \llbracket \Delta, u : B, \Delta'\,;\Gamma \rrbracket \to \llbracket \Delta, \Delta'\,;\Gamma \rrbracket$ *is the obvious projection.*

Proof. By induction on the two derivations. All cases are straightforward. The modal one uses Propositions 5.11(2), 5.17(5), 5.22(2), and 5.33. □

Lemma 6.6 (Semantics of Exchange)**.**
(1) *Let* $\Delta\,;\Gamma, x : C, y : D, \Gamma' \vdash M : A$. *Then*
$$\llbracket \Delta\,;\Gamma, x : C, y : D, \Gamma' \vdash M : A \rrbracket = \llbracket \Delta\,;\Gamma, y : D, x : C, \Gamma' \vdash M : A \rrbracket \circ (\cong)$$
where $(\cong) : \llbracket \Delta\,;\Gamma, x : C, y : D, \Gamma' \rrbracket \xrightarrow{\cong} \llbracket \Delta\,;\Gamma, y : D, x : C, \Gamma' \rrbracket$ *is the obvious isomorphism.*
(2) *Let* $\Delta, u : C, v : D, \Delta'\,;\Gamma \vdash M : A$. *Then*
$$\llbracket \Delta, u : C, v : D, \Delta'\,;\Gamma \vdash M : A \rrbracket = \llbracket \Delta, v : D, u : C, \Delta'\,;\Gamma \vdash M : A \rrbracket \circ (\cong)$$
where $(\cong) : \llbracket \Delta, u : C, v : D\,;\Gamma \rrbracket \xrightarrow{\cong} \llbracket \Delta, v : D, u : C\,;\Gamma \rrbracket$ *is the obvious isomorphism.*

Proof. By induction on the two derivations. All cases are straightforward. □

Then, we move on to something particular to the cases of T and $\mathsf{S4}$, namely the interpretation of the Modal Dereliction rule—see Theorem 3.9.

Lemma 6.7 (Semantics of Dereliction)**.** *Let* $\Delta\,;\Gamma,\Gamma' \vdash_{D\mathcal{L}} M : A$ *where* $\mathcal{L} \in \{\mathsf{T}, \mathsf{S4}\}$ *and* $\Gamma = \vec{z} : \vec{C}$*. Then*

$$\llbracket \Delta, \Gamma\,; \Gamma' \vdash M : A \rrbracket_{\mathcal{L}} = \llbracket \Delta\,; \Gamma, \Gamma' \vdash M : A \rrbracket_{\mathcal{L}} \circ \left(\overrightarrow{id_{\Delta}} \times \overrightarrow{\epsilon_{C_i}} \times \overrightarrow{id_{\Gamma'}} \right)$$

Proof. By induction on the derivation of $\Delta\,;\Gamma,\Gamma' \vdash_{\mathsf{D}\mathcal{L}} M : A$. All cases are straightforward. The case for $(\Box\mathcal{E})$ depends on the semantics of exchange lemma. $\square$

We also need to know that 'boxing' a variable results in the obvious projection. This depends essentially on the fact our functors are product-preserving (and not just lax monoidal).

Lemma 6.8 (Identity Lemma)**.** *For* $(u_i : B_i) \in \Delta$*, and* $\mathcal{L} \in \{\mathsf{K}, \mathsf{K4}, \mathsf{T}, \mathsf{S4}\}$*,*

$$\llbracket \Delta\,; \Gamma \vdash \mathsf{box}\ u_i : \Box B_i \rrbracket_{\mathcal{L}} = \pi^{\Delta;\Gamma}_{\Box B_i}$$

We aim to show that substitution in the syntax corresponds to composition in the semantics. To make this result work, we need to introduce a $\mathsf{box}\,(-)$ construct for GL. We write

$$\mathsf{box}\ M \stackrel{\text{def}}{=} \mathsf{fix}\ w\ \mathsf{in}\ \mathsf{box}\ M$$

with $w, w^{\perp}$ fresh. It is then not hard to see that the introduction rule of $\mathsf{K4}$ is admissible for GL when M has no occurrences of $w^{\perp}$: we simply use weakening followed by the introduction rule for GL. This derived operation is reflected in the semantics by the equation

Proposition 6.9. $\llbracket \Delta\,; \Gamma \vdash_{DGL} \mathsf{box}\ M : \Box A \rrbracket = \llbracket \Delta\,; \Delta^{\perp} \vdash_{DGL} M^{\perp} : A \rrbracket^{\#} \circ \pi^{\Delta;\Gamma}_{\Delta}$

Proof. By the semantics of weakening and Proposition 5.35. $\square$

In short: when the variable that is being 'diagonalised over' does *not* occur freely, the interpretation degenerates to that of $\mathsf{K4}$.

Lemma 6.10 (Semantics of Substitution)**.** *Suppose that* $\vec{u} : \vec{B}\,; \vec{x} : \vec{A} \vdash_{\mathsf{D}\mathcal{L}} P : C$*. Let* $\Delta\,; \Gamma \vdash_{\mathsf{D}\mathcal{L}} M_i : A_i$ *for* $i = 1, \ldots, n$*, and let*

$$\alpha_i \stackrel{\text{def}}{=} \llbracket \Delta\,; \Gamma \vdash M_i : A_i \rrbracket_{\mathcal{L}}$$

If either

(1) $\mathcal{L} \in \{\mathsf{K}, \mathsf{T}\}$ *and* $\cdot\,; \Delta \vdash N_j : B_j$ *for* $j = 1, \ldots, m$*, or*
(2) $\mathcal{L} \in \{\mathsf{K4}, \mathsf{GL}\}$ *and* $\Delta\,; \Delta^{\perp} \vdash N_j^{\perp} : B_j$ *for* $j = 1, \ldots, m$*, or*
(3) $\mathcal{L} = \mathsf{S4}$ *and* $\Delta\,; \cdot \vdash N_j : B_j$ *for* $j = 1, \ldots, m$*,*

then, letting $\beta_j \stackrel{\text{def}}{=} \llbracket \Delta\,; \Gamma \vdash \mathsf{box}\ N_j : \Box B_j \rrbracket_{\mathcal{L}}$ *for* $j \in \{1, \ldots, m\}$*, we have that*

$$\left\llbracket \Delta\,; \Gamma \vdash P[\vec{N}/\vec{u}, \vec{M}/\vec{x}] : C \right\rrbracket_{\mathcal{L}} = \left\llbracket \vec{u} : \vec{B}\,; \vec{x} : \vec{A} \vdash P : C \right\rrbracket_{\mathcal{L}} \circ \langle \beta_1, \ldots, \beta_m, \alpha_1, \ldots, \alpha_n \rangle$$

Proof. By induction on the derivation of $\vec{u} : \vec{B}; \vec{x} : \vec{A} \vdash P : C$. Most cases are straightforward, and use a combination of standard equations that hold in cartesian closed categories in order to perform calculations very close the ones detailed in [AT11, §1.6.5]. Because of the precise definitions we have used, we also need to make use of Lemma 6.5 to interpret weakening whenever variables in the context do not occur freely in the term. For the modal rules we use many of the equations we showed in §5, e.g in Propositions 5.11, 5.17, 5.22, and so on. $\square$

Theorem 6.11 (Soundness)**.** *If $\Delta\,;\Gamma \vdash_{\mathsf{D}\mathcal{L}} M = N : A$, then we have that*

$$[\![\Delta\,;\Gamma \vdash M : A]\!]_{\mathcal{L}} = [\![\Delta\,;\Gamma \vdash N : A]\!]_{\mathcal{L}}$$

Proof. By induction on the derivation of $\Delta\,;\Gamma \vdash_{\mathsf{D}\mathcal{L}} M = N : A$. The congruence cases are clear, as are the majority of the ordinary clauses—see [Cro93] and [AT11]. The rules that remain are $(\Box\eta)$, the many variants of $(\Box\beta)$, and the commuting conversions.

First, we prove the modal β and η cases by direct calculation. To do so, we use Lemma 6.8, so product preservation is essential even to prove the soundness of $(\Box\beta)$.

Let $\Delta = \vec{u} : \vec{B}$ and $\Gamma = \vec{x} : \vec{A}$. We then calculate:

$$\begin{aligned}
& [\![\Delta\,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow \mathsf{box}\ M\ \mathsf{in}\ N : C]\!] \\
= \;& \{\,\text{definition}\,\} \\
& [\![\Delta, u : A\,;\Gamma \vdash N : C]\!] \circ \langle \overrightarrow{\pi_\Delta}, [\![\Delta\,;\Gamma \vdash \mathsf{box}\ M : \Box A]\!], \overrightarrow{\pi_\Gamma}\rangle \\
= \;& \{\,\text{Lemma 6.8}\,\} \\
& [\![\Delta, u : A\,;\Gamma \vdash N : C]\!] \circ \langle \overrightarrow{[\![\Delta\,;\Gamma \vdash \mathsf{box}\ u_i : \Box B_i]\!]}, [\![\Delta\,;\Gamma \vdash \mathsf{box}\ M : \Box A]\!], \overrightarrow{[\![\Delta\,;\Gamma \vdash x_i : A_i]\!]}\rangle \\
= \;& \{\,\text{Lemma 6.10}\,\} \\
& [\![\Delta\,;\Gamma \vdash N[\vec{u_i}/\vec{u_i}, M/u, \vec{x_i}/\vec{x_i}] : C]\!]
\end{aligned}$$

This covers all cases save GL. For that, it suffices to show that

$$[\![\Delta\,;\Gamma \vdash \mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M : \Box A]\!] = [\![\Delta\,;\Gamma \vdash \mathsf{box}\ M[\mathsf{fix}\ z\ \mathsf{in}\ \mathsf{box}\ M/z] : \Box A]\!]$$

and then surreptitiously swap the first expression with the second in the above calculation just before using the substitution lemma.

The case of η is even simpler, as it follows immediately from Lemma 6.8.

The commuting conversions for weakening and contraction are straightforward. $(\mathsf{commlet})$ requires a nested induction on contexts $C[-]$, which follows from the naturality of the various operations of the CCC. $\Box$

Idempotence. If the comonad (F, ϵ, δ) provided as part of a Bierman-de Paiva category is idempotent, then more equations are sound. We have shown the equivalence between three such equations in §6.1, so it suffices to prove soundness for only one of them:

$$\begin{aligned}
& [\![\Delta\,;\Gamma \vdash \mathsf{box}\ (\mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in}\ N) : \Box B]\!] \\
= \;& \{\ \text{definitions}\ \} \\
& ([\![\Delta, u : A\,;\cdot \vdash N : B]\!] \circ \langle id, [\![\Delta\,;\cdot \vdash M : \Box A]\!]\rangle)^* \\
= \;& \{\ \text{Proposition Theorem 5.19}\ \} \\
& [\![\Delta, u : A\,;\cdot \vdash N : B]\!]^* \circ \langle id, [\![\Delta\,;\cdot \vdash M : \Box A]\!]\rangle \\
= \;& \{\ \text{definitions}\ \} \\
& [\![\Delta\,;\Gamma \vdash \mathsf{let\ box}\ u \Leftarrow M\ \mathsf{in\ box}\ N : \Box B]\!]
\end{aligned}$$

6.4. **A note on completeness.** It is possible to prove that the categorical semantics given in this section are *complete* in the Lindenbaum-Tarski sense. For example, we can prove that if $[\![\Delta\,;\Gamma \vdash_{\mathsf{DK}} M : A]\!] = [\![\Delta\,;\Gamma \vdash_{\mathsf{DK}} N : A]\!]$ in every Kripke category, then $\Delta\,;\Gamma \vdash_{\mathsf{DK}} M = N : A$ is provable in the equational theory. To do so we must construct a Kripke category by quotienting the syntax of the λ-calculus; if the equation is satisfied in all models, it is satisfied in this one in particular, which implies equality in the theory.

Indeed, we have shown this for all our calculi, as documented in the conference version of this article and its associated technical report [Kav17]. However, this result is of limited interest from the point-of-view of categorical logic: it merely shows that the correspondence between inference rules and axioms (e.g. between Scott's rule and the K axiom) extends to the categorical level, e.g. to a correspondence between the operation $(-)^\bullet$ and the product-preserving structure of a Kripke category.

A more important result is to show that the model constructed through quotienting syntax is *initial* in the category of all such models. Unfortunately, it does not seem that the constructions used in the conference version of this paper yield initial models. We leave the solution of this problem to future work.

7. Conclusion & further directions

We have extended the full Curry-Howard-Lambek correspondence to a handful of normal modal logics, spanning the logical aspect (Hilbert systems and provability), the computational aspect (modal λ-calculi), and the categorical aspect (proof-relevant semantics).

In order to achieve the connection at the first junction, i.e. that between logic and computation, we have employed a systematic pattern based on translating sequent calculus rules to introduction rules for dual-context systems. This worked remarkably well: not only did it lead to already known systems, like that for $\mathsf{S4}$, but also to a number of new ones, including one for GL. One would hope that there is a deeper aspect to this pattern—perhaps even a theorem to the effect that rules of cut-free sequent calculi rule can be immediately turned into well-behaved dual-context systems. Of course, this is quite a long way from our current grasp, but we believe it is worth investigating.

The second junction, i.e. the one between modal λ-calculi and their categorical models, is achieved by focusing on product-preserving functors and their extensions. This deviates from previous approaches—such as that of Bierman and de Paiva which concentrated on lax monoidal endofunctors—yet supports numerous models of interest. We sketched in some detail the topos of bifurcating trees as a motivating example of a model with a product-preserving modality which does not extend to a comonad. The assumption of product preservation seeps deep into our proofs: it is in fact used even in proving soundness of β-convertibility.

The resulting dual-context calculi sport a simpler syntax, which—as we argued in the introduction—makes them particularly suited to computational applications. It is also our hope that this work will help elucidate the computational behaviour of necessity modalities. In fact, the author believes that modalities can be used to control the 'flow of data' in a programming language, in the sense that they create regions of the language whose intercommunication is restricted. For example, one can handwavingly argue that $\mathsf{S4}$ guarantees that 'only modal variables flow into terms of modal type,' whereas K additionally ensures that no modal data flows into a term of non-modal type. A first result of this type is the free variables theorem (Theorem 3.5), but it is rather weak. The author has recently used the second junction—namely that between computation and categories—to prove *noninterference theorems* for modal λ-calculi [Kav19]. These theorems show that modal type systems indeed effect certain restrictions on information flow. Amongst other things, the present article is meant to lay a foundation that enables the further study of categorical semantics of similar calculi.

Acknowledgments

I gratefully acknowledge many interesting discussions with Daniel Gratzer, Dan Licata, Kristina Sojakova, Mitchell Riley, and Ed Morehouse. Thanks are also due to Sam Staton and Luke Ong for their encouragement, Samson Abramsky for his advice, and Geraint Jones for his Eindhoven-style calculation macros.

The lion's share of this work was done during the author's time as a doctoral student at Oxford, where he was supported by the EPSRC (award reference 1354534). It was then revised at Wesleyan University, as part of work supported by the Air Force Office of Scientific Research under award number FA9550-16-1-0292. Any opinions, finding, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Air Force. Its completion was supported in part by a research grant (12386, Guarded Homotopy Type Theory) from the VILLUM Foundation.

References

[AB04] Steven Awodey and Andrej Bauer. Propositions as [Types]. *Journal of Logic and Computation*, 14(4):447–471, aug 2004.

[ABHR99] Martín Abadi, Anindya Banerjee, Nevin Heintze, and Jon G Riecke. A core calculus of dependency. In *Proceedings of the 26th ACM SIGPLAN-SIGACT symposium on Principles of programming languages - POPL '99*, pages 147–160, New York, New York, USA, 1999. ACM Press.

[AM13] Robert Atkey and Conor McBride. Productive coprogramming with guarded recursion. *ACM SIGPLAN Notices*, 48(9):197–208, nov 2013.

[And92] Jean-Marc Andreoli. Logic Programming with Focusing Proofs in Linear Logic. *Journal of Logic and Computation*, 2(3):297–347, 1992.

[AT11] Samson Abramsky and Nikos Tzevelekos. Introduction to Categories and Categorical Logic. In Bob Coecke, editor, *New Structures for Physics*, pages 3–94. Springer-Verlag, 2011.

[Awo10] Steve Awodey. *Category Theory*. Oxford Logic Guides. Oxford University Press, 2010.

[Bar96] Andrew Graham Barber. Dual Intuitionistic Linear Logic. Technical report, ECS-LFCS-96-347, Laboratory for Foundations of Computer Science, University of Edinburgh, 1996.

[BBdP98] Nick Benton, Gavin M. Bierman, and Valeria de Paiva. Computational types from a logical perspective. *Journal of Functional Programming*, 8(2):177–193, 1998.

[BBdPH93] Nick Benton, Gavin Bierman, Valeria de Paiva, and Martin Hyland. A term calculus for Intuitionistic Linear Logic. In Marc Bezem and Jan Friso Groote, editors, *Typed Lambda Calculi and Applications*, pages 75–90, Berlin, Heidelberg, 1993. Springer Berlin Heidelberg.

[BdP92] Gavin M. Bierman and Valeria de Paiva. Intuitionistic Necessity Revisited. In *Proceedings of the Logic at Work Conference*, 1992.

[BdP96] Gavin M. Bierman and Valeria de Paiva. Intuitionistic Necessity Revisited. Technical report, University of Birmingham, 1996.

[BdP00] Gavin M. Bierman and Valeria de Paiva. On an Intuitionistic Modal Logic. *Studia Logica*, 65(3):383–416, 2000.

[BdPR01] Gianluigi Bellin, Valeria de Paiva, and Eike Ritter. Extended Curry-Howard correspondence for a basic constructive modal logic. In *Proceedings of Methods for Modalities*, 2001.

[BdRV01] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*. Cambridge University Press, 2001.

[Bel85] Gianluigi Bellin. A system of natural deduction for GL. *Theoria*, 51(2):89–114, 1985.

[BGM19] Patrick Bahr, Christian Uldal Graulund, and Rasmus Ejlers Møgelberg. Simply RaTT: a fitch-style modal calculus for reactive programming without space leaks. *Proceedings of the ACM on Programming Languages*, 3(ICFP):1–27, 2019.

[BM13] Lars Birkedal and Rasmus Ejlers Møgelberg. Intensional Type Theory with Guarded Recursive Types qua Fixed Points on Universes. In *2013 28th Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 213–222. IEEE, jun 2013.

[BMSS12] Lars Birkedal, Rasmus Møgelberg, Jan Schwinghammer, and Kristian Støvring. First steps in synthetic guarded domain theory: step-indexing in the topos of trees. *Logical Methods in Computer Science*, 8(4), oct 2012.

[Boo94] George S. Boolos. *The Logic of Provability*. Cambridge University Press, Cambridge, feb 1994.

[Bor94] Francis Borceux. *Handbook of Categorical Algebra*. Cambridge University Press, Cambridge, 1994.

[CBBB16] Ranald Clouston, Alĕs Bizjak, Hans Bugge Grathwohl, and Lars Birkedal. The guarded lambda calculus: Programming and reasoning with guarded recursion for coinductive types. *Logical Methods in Computer Science*, 12(3):1–39, 2016.

[CF58] Haskell B. Curry and Robert Feys. *Combinatory Logic*. Studies in Logic and the Foundation of Mathematics. North-Holland, 1958.

[CFMM16] Pierre-Louis Curien, Marcelo Fiore, and Guillaume Munch-Maccagnoni. A theory of effects and resources: adjunction models and polarised calculi. In *Proceedings of the 43rd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages - POPL 2016*, pages 44–56, New York, New York, USA, 2016. ACM Press.

[Clo18] Ranald Clouston. Fitch-style modal lambda calculi. In Christel Baier and Ugo Dal Lago, editors, *Foundations of Software Science and Computation Structures*, pages 258–275, Cham, 2018. Springer International Publishing.

[Cro93] Roy L. Crole. *Categories for Types*. Cambridge University Press, 1993.

[Cur52] Haskell B. Curry. The elimination theorem when modality is present. *The Journal of Symbolic Logic*, 17(04):249–265, dec 1952.

[Dav17] Rowan Davies. A Temporal Logic Approach to Binding-Time Analysis. *Journal of the ACM*, 64(1):1–45, mar 2017.

[dG02] Philippe de Groote. On the Strong Normalisation of Intuitionistic Natural Deduction with Permutation-Conversions. *Information and Computation*, 178(2):441–464, 2002.

[DJ03] Vincent Danos and Jean Baptiste Joinet. Linear logic and elementary time. *Information and Computation*, 183(1):123–137, 2003.

[DP01] Rowan Davies and Frank Pfenning. A modal analysis of staged computation. *Journal of the ACM*, 48(3):555–604, 2001.

[dPR11] Valeria de Paiva and Eike Ritter. Basic Constructive Modality. In Jean-Yves Beziau and Marcelo Coniglio, editors, *Logic without frontiers - Festschrift for Walter Alexandre Carnielli on the occasion of his 60th birthday*, pages 411–428. College Publications, London, 2011.

[Fit93] Melvin Fitting. *Basic Modal Logic*, pages 368–448. Oxford University Press, Inc., USA, 1993.

[Gal90] Jean Gallier. On Girard's "Candidats de Reductibilite". In Piergiorgio Odifreddi, editor, *Logic and Computer Science*, pages 123–203. Academic Press, 1990.

[Gal93] Jean Gallier. Constructive logics Part I: A tutorial on proof systems and typed λ-calculi. *Theoretical Computer Science*, 110(2):249–339, 1993.

[Gal95] Jean Gallier. On the Correspondence Between Proofs and Lambda Terms. In Philippe de Groote, editor, *The Curry-Howard Isomorphism*, pages 55–138. Academia, Louvain-la-Neuve, 1995.

[Gen35a] Gerhard Gentzen. Untersuchungen über das logische Schließen. I. *Mathematische Zeitschrift*, 39(1):176–210, 1935.

[Gen35b] Gerhard Gentzen. Untersuchungen über das logische Schließen. II. *Mathematische Zeitschrift*, 39(1):405–431, 1935.

[Gir72] Jean-Yves Girard. *Interprétation fonctionelle et élimination des coupures de l'arithmétique d'ordre supérieur*. PhD thesis, Université Paris VII, 1972.

[Gir93] Jean-Yves Girard. On the unity of logic. *Annals of Pure and Applied Logic*, 59(3):201–217, feb 1993.

[GL96] Jean Goubault-Larrecq. On Computational Interpretations of the Modal Logic S4 - I. Cut Elimination. Technical report, 1996-35. Institut für Logik, Komplexität und Deduktionssysteme, Universität Karlsruhe, 1996.

[GLN+17] Jacob A Gross, Daniel R Licata, Max S New, Jennifer Paykin, Mitchell Riley, Michael Shulman, and Felix Wellen. Differential Cohesive Type Theory (Extended Abstract). In *Extended abstracts for the Workshop "Homotopy Type Theory and Univalent Foundations"*, 2017.

[GLT89] Jean-Yves Girard, Yves Lafont, and Paul Taylor. *Proofs and Types*. Cambridge University Press, 1989.

[GR12] Rajeev Goré and Revantha Ramanayke. Valentini's Cut-Elimination for Provability Logic Resolved. *The Review of Symbolic Logic*, 5(02):212–238, 2012.

[Gua18] Adrien Guatto. A Generalized Modality for Recursion. In *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '18, pages 482–491, New York, NY, USA, 2018. Association for Computing Machinery.

[HC96] G. E. Hughes and M. J. Creswell. *A New Introduction to Modal Logic*. Routledge, 1996.

[HJ99] K Hrbacek and T Jech. *Introduction to Set Theory, Third Edition, Revised and Expanded*. Chapman & Hall/CRC Pure and Applied Mathematics. Taylor & Francis, 1999.

[HN12] Raul Hakli and Sara Negri. Does the deduction theorem fail for modal logic? *Synthese*, 187(3):849–867, 2012.

[Hof99a] Martin Hofmann. Semantical analysis of higher-order abstract syntax. In *Proceedings. 14th Symposium on Logic in Computer Science (Cat. No. PR00158)*, pages 204–213. IEEE Comput. Soc, 1999.

[Hof99b] Martin Hofmann. *Type Systems for Polynomial-Time Computation*. Habilitation thesis, Technischen Universität Darmstadt, 1999.

[Kak07] Yoshihiko Kakutani. Calculi for Intuitionistic Normal Modal Logic. In *Proceedings of Programming and Programming Languages (PPL) 2007*, 2007.

[Kav16] G. A. Kavvos. The Many Worlds of Modal Lambda Calculi: I. Curry-Howard for Necessity, Possibility and Time. *CoRR*, 2016.

[Kav17] G. A. Kavvos. Dual-context calculi for modal logic. In *2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE, 2017.

[Kav19] G. A. Kavvos. Modalities, Cohesion, and Information Flow. *Proceedings of the ACM on Programming Languages*, 3(POPL), 2019.

[KB11a] Neelakantan R. Krishnaswami and Nick Benton. A semantic model for graphical user interfaces. *ACM SIGPLAN Notices*, 46(9):45, sep 2011.

[KB11b] Neelakantan R. Krishnaswami and Nick Benton. Ultrametric Semantics of Reactive Programs. In *2011 IEEE 26th Annual Symposium on Logic in Computer Science*, pages 257–266. IEEE, jun 2011.

[Kol85] George Koletsos. Church-Rosser theorem for typed functional systems. *The Journal of Symbolic Logic*, 50(03):782–790, 1985.

[Kri63] Saul A. Kripke. Semantical Analysis of Modal Logic I. Normal Modal Propositional Calculi. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 9(5-6):67–96, 1963.

[Kri13] Neelakantan R. Krishnaswami. Higher-order functional reactive programming without space-time leaks. In *Proceedings of the 18th ACM SIGPLAN international conference on Functional programming - ICFP '13*, page 221, New York, New York, USA, 2013. ACM, ACM Press.

[Lei81] Daniel Leivant. On the proof theory of the modal logic for arithmetic provability. *The Journal of Symbolic Logic*, 46(03):531–538, sep 1981.

[Lit14] Tadeusz Litak. Constructive Modalities with Provability Smack. In G. Bezhanishvili, editor, *Leo Esakia on Duality in Modal and Intuitionistic Logics*, volume 4 of *Outstanding Contributions to Logic*, pages 187–216. Springer, Dordrecht, 2014.

[LOPS18] Daniel R. Licata, Ian Orton, Andrew M. Pitts, and Bas Spitters. Internal Universes in Models of Homotopy Type Theory. In H. Kirchner, editor, *3rd International Conference on Formal Structures for Computation and Deduction (FSCD 2018)*, Leibniz International Proceedings in Informatics (LIPIcs), pages 22:1–22:17. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

[LS88] Joachim Lambek and Philip J. Scott. *Introduction to Higher-Order Categorical Logic*. Cambridge University Press, 1988.

[Mac78] Saunders Mac Lane. *Categories for the Working Mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer New York, New York, NY, 1978.

[MCHP04] T. Murphy, Karl Crary, Robert Harper, and Frank Pfenning. A symmetric modal lambda calculus for distributed computing. In *Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science, 2004.*, pages 286–295. IEEE, 2004.

[Mel09] Paul-André Melliès. Categorical Semantics of Linear Logic. In Pierre-Louis Curien, Hugo Herbelin, Jean-Louis Krivine, and Paul-André Melliès, editors, *Panoramas et synthèses 27: Interactive models of computation and program behaviour*. Société Mathématique de France, 2009.

[ML13] Stefan Milius and Tadeusz Litak. Guard Your Daggers and Traces: On The Equational Properties of Guarded (Co-)recursion. *Electronic Proceedings in Theoretical Computer Science*, 126(Informatik 8):72–86, aug 2013.

[MM96] Simone Martini and Andrea Masini. A Computational Interpretation of Modal Proofs. In *Proof Theory of Modal Logic*, Applied Logic Series, pages 213–241. Springer Netherlands, 1996.

[Mog91] Eugenio Moggi. Notions of computation and monads. *Information and Computation*, 93(1):55–92, 1991.

[Nak00] Hiroshi Nakano. A modality for recursion. *Proceedings Fifteenth Annual IEEE Symposium on Logic in Computer Science (Cat. No.99CB36332)*, 2000.

[ND18] Andreas Nuyts and Dominique Devriese. Degrees of Relatedness. In *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science - LICS '18*, pages 779–788, New York, New York, USA, 2018. ACM Press.

[Neg11] Sara Negri. Proof Theory for Modal Logic. *Philosophy Compass*, 6(8):523–538, 2011.

[NVD17] Andreas Nuyts, Andrea Vezzosi, and Dominique Devriese. Parametric quantifiers for dependent type theory. *Proceedings of the ACM on Programming Languages*, 1(ICFP), 2017.

[OH06] Yo Ohta and Masahito Hasegawa. A Terminating and Confluent Linear Lambda Calculus. In Frank Pfenning, editor, *Term Rewriting and Applications. RTA 2006*, volume 4098 of *Lecture Notes in Computer Science*, pages 166–180. Springer, Berlin, Heidelberg, 2006.

[OM57] Masao Ohnisi and Kazuo Matsumoto. Gentzen method in modal calculi. *Osaka Journal of Mathematics*, 11(2):113–130, 1957.

[OM59] Masao Ohnisi and Kazuo Matsumoto. Gentzen method in modal calculi. II. *Osaka Journal of Mathematics*, 11(2):115–120, 1959.

[Ono98] Hiroakira Ono. Proof-theoretic methods in nonclassical logic–an introduction. In Masako Takahashi, Mitsuhiro Okada, and Marangiola Dezani-Ciancaglini, editors, *Theories of Types and Proofs*, MSJ Memoirs, pages 207–254. The Mathematical Society of Japan, Tokyo, 1998.

[Orc14] Dominic Orchard. *Programming contextual computations.* PhD thesis, University of Cambridge, 2014.

[PD01] Frank Pfenning and Rowan Davies. A judgmental reconstruction of modal logic. *Mathematical Structures in Computer Science*, 11(4):511–540, 2001.

[Pfe01] Frank Pfenning. Intensionality, extensionality, and proof irrelevance in modal type theory. *Proceedings of the 16th Annual IEEE Symposium on Logic in Computer Science (LICS 2001)*, 2001.

[Pfe10] Frank Pfenning. Lecture Notes on Combinatory Modal Logic, 2010.

[Pfe13] Frank Pfenning. Weather Report, 2013.

[Pfe15] Frank Pfenning. Decomposing Modalities. *Logical Frameworks and Meta-Languages: Theory and Practice (LFMTP'15)*, 2015.

[Plo77] Gordon D. Plotkin. LCF considered as a programming language. *Theoretical Computer Science*, 5(3):223–255, 1977.

[Plo93] Gordon D. Plotkin. Type theory and recursion. In *Proceedings Eighth Annual IEEE Symposium on Logic in Computer Science*, page 374. IEEE Comput. Soc. Press, 1993.

[Pra65] Dag Prawitz. *Natural Deduction: a proof-theoretical study.* Almquist and Wiksell, 1965.

[Pra71] Dag Prawitz. Ideas and Results in Proof Theory. In J. E. Fenstad, editor, *Proceedings of the Second Scandinavian Logic Symposium*, volume 63 of *Studies in logic and the foundations of mathematics*. North-Holland, Amsterdam, 1971.

[RSS20] Egbert Rijke, Michael Shulman, and Bas Spitters. Modalities in homotopy type theory. *Logical Methods in Computer Science*, 16(1):2:1–2:79, 2020.

[Sch17] Gabriel Scherer. Deciding equivalence with sums and the empty type. In *Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages - POPL 2017*, pages 374–386, New York, New York, USA, 2017. ACM Press.

[SDP01] Carsten Schürmann, Joëlle Despeyroux, and Frank Pfenning. Primitive recursion for higher-order abstract syntax. *Theoretical Computer Science*, 266(1-2):1–57, 2001.

[SdV12] Paula G. Severi and Fer-Jan J. de Vries. Pure type systems with corecursion on streams. *ACM SIGPLAN Notices*, 47(9):141, 2012.

[SH84] Peter Schroeder-Heister. A natural extension of natural deduction. *The Journal of Symbolic Logic*, 49(04):1284–1300, 1984.

[Sha14] D. S. Shamkanov. Circular proofs for the Gödel-Löb provability logic. *Mathematical Notes*, 96(3-4):575–585, sep 2014.

[Shu18] Michael Shulman. Brouwer's fixed-point theorem in real-cohesive homotopy type theory. *Mathematical Structures in Computer Science*, 28(6):856–941, 2018.

[SI08] Naokata Shikuma and Atsushi Igarashi. Proving Noninterference by a Fully Complete Translation to the Simply Typed lambda-calculus. *Logical Methods in Computer Science*, 4(3):10, 2008.

[Sim94] Alex K. Simpson. *The Proof Theory and Semantics of Intuitionistic Modal Logic*. PhD thesis, The University of Edinburgh, 1994.

[SP00] Alex K. Simpson and Gordon D. Plotkin. Complete axioms for categorical fixed-point operators. In *Proceedings of the 15th Annual IEEE Symposium on Logic in Computer Science (LICS 2000)*, pages 30–41. IEEE Comput. Soc, 2000.

[SS14] Urs Schreiber and Michael Shulman. Quantum Gauge Field Theory in Cohesive Homotopy Type Theory. *Electronic Proceedings in Theoretical Computer Science*, 158:109–126, 2014.

[SU06] Morten Heine Sørensen and Pawel Urzyczyn. *Lectures on the Curry-Howard Isomorphism*. Elsevier, 2006.

[SV80] G. Sambin and S. Valentini. A modal sequent calculus for a fragment of arithmetic. *Studia Logica*, 39(2-3):245–256, 1980.

[SV82] Giovanni Sambin and Silvio Valentini. The modal logic of provability. The sequential approach. *Journal of Philosophical Logic*, 11(3):311–342, 1982.

[Tak95] M. Takahashi. Parallel Reductions in λ-Calculus. *Information and Computation*, 118(1):120–127, apr 1995.

[TI10] Takeshi Tsukada and Atsushi Igarashi. A logical foundation for environment classifiers. *Logical Methods in Computer Science*, 6(4):1–43, 2010.

[Val82] Silvio Valentini. Cut-elimination in a modal sequent calculus for K. *Bolletino dell'Unione Mathematica Italiana*, 1B:119–130, 1982.

[Val83] Silvio Valentini. The Modal Logic of Provability: Cut-Elimination. *Journal of Philosophical Logic*, 12(4):471–476, 1983.

[Wad93] Philip Wadler. A taste of linear logic. In Andrzej M Borzyszkowski and Stefan Sokołowski, editors, *Mathematical Foundations of Computer Science 1993*, pages 185–210, Berlin, Heidelberg, 1993. Springer Berlin Heidelberg.

[Wad94] Philip Wadler. A syntax for linear logic. In S. Brookes, M. Main, A. Melton, M. Mislove, and D. Schmidt, editors, *Mathematical Foundations of Programming Semantics: 9th International Conference, New Orleans, LA, USA, April 7 - 10, 1993. Proceedings*, pages 513–529. Springer-Verlag Berlin Heidelberg, 1994.

[Wan02] Heinrich Wansing. Sequent Systems for Modal Logics. In *Handbook of Philosophical Logic*, pages 61–145. Springer Netherlands, Dordrecht, 2002.

[Wel17] Felix Wellen. *Formalizing Cartan Geometry in Modal Homotopy Type Theory*. Phd thesis, Karlsruher Instituts für Technologie, 2017.

[Wij90] Duminda Wijesekera. Constructive modal logics I. *Annals of Pure and Applied Logic*, 50(3):271–301, dec 1990.

[WLP98] Philip Wickline, Peter Lee, and Frank Pfenning. Run-Time Code Generation and Modal-ML. In *Proceedings of the ACM SIGPLAN 1998 Conference on Programming Language Design and Implementation*, PLDI '98, pages 224–235, New York, NY, USA, 1998. Association for Computing Machinery.

This work is licensed under the Creative Commons Attribution License. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/ or send a letter to Creative Commons, 171 Second St, Suite 300, San Francisco, CA 94105, USA, or Eisenacher Strasse 2, 10777 Berlin, Germany