MAX-PLANCK-INSTITUT FÜR INFORMATIK

Labelled Propositional Modal Logics:
Theory and Practice

David Basin
Seán Matthews
Luca Viganò

MPI–I–96–2–002 February 1996

Im Stadtwald
D 66123 Saarbrücken
Germany

Authors' Addresses

David Basin, Seán Matthews, Luca Viganò
Max-Planck-Institut für Informatik, Im Stadtwald, D-66123 Saarbrücken, Germany
{basin,sean,luca}mpi-sb.mpg.de

Publication Notes

The present report has been submitted for publication elsewhere and will be copyrighted if accepted.

Abstract

We show how labelled deductive systems can be combined with a logical framework to provide a natural deduction implementation of a large and well-known class of propositional modal logics (including K, D, T, B, $S4$, $S4.2$, $KD45$, $S5$). Our approach is modular and based on a separation between a base logic and a labelling algebra, which interact through a fixed interface. While the base logic stays fixed, different modal logics are generated by plugging in appropriate algebras. This leads to a hierarchical structuring of modal logics with inheritance of theorems. Moreover, it allows modular correctness proofs, both with respect to soundness and completeness for semantics, and faithfulness and adequacy of the implementation. We also investigate the tradeoffs in possible labelled presentations: We show that a narrow interface between the base logic and the labelling algebra supports modularity and provides an attractive proof-theory (in comparision to, e.g., semantic embedding) but limits the degree to which we can make use of extensions to the labelling algebra.

Keywords

Labelled Deductive Systems, Modal Logics, Consequence Relations, Logical Frameworks, Combination of Logics

Contents

1 Introduction

In this paper we examine how two complementary proposals for dealing with the enormous range of logics developed in recent years combine together in practice. The first is the use of a generic theorem prover [10, 11, 15], based on a logical framework, which can be used to implement proof systems for many logics in a uniform manner. These theorem provers are based on a metalogic in which the syntax and proof rules of object logics are encoded, and theorems of the object logic are constructed by proving theorems in the metalogic. The second is that of a Labelled Deductive System (LDS, [8]), a method for giving uniform presentations of non-standard logics based on possibly radically different deductive systems, e.g. modal, substructural, or non-monotonic logics. In the LDS approach, instead of a consequence relation being defined over formulae ($\ldots A \vdash B \ldots$), it is defined over pairs consisting of a label and a formula ($\ldots x : A \vdash y : B \ldots$). The labels then allow information needed to formalize the more subtle metatheoretic aspects of the relation to be tracked. For modal logic, for instance, we might want to distinguish between 'local' (with respect to some world) and 'global' (with respect to some frame) consequence, so the label could keep track of the 'possible world' in which the formula lives. Or for a substructural logic, where the consequence relation should be sensitive to operations like weakening and contraction, the labels might track resources and their use [5].

We study this combination in the case of propositional modal logics and show how it can provide a simple and usable implementation of a large collection of logics (including K, D, T, B, $S4$, $S4.2$, $KD45$, $S5$) in a natural deduction (ND, [16, 17]) setting. We view a proof system for an LDS as consisting of two parts: a *base logic* for manipulating labelled formulae, and a separate *labelling algebra* for reasoning about the labels. Our base logic, in which labels represent possible worlds in the Kripke frame, is a labelled ND presentation of propositional calculus extended with introduction and elimination rules for $\Box$ (i.e. the modal logic K). Our labelling algebras are *relational theories* comprised of Horn clause axioms formalizing the accessibility of worlds in Kripke frames. These two parts are separate and communicate through an *interface* provided by the rules for $\Box$. We implement these theories in the Isabelle logical framework [15], and this separation is enforced by the use of multiple judgements (cf. [10]) in the metalogic, which distinguish between relational and labelled formulae.

Why Combine Paradigms?

Why should the LDS and logical framework paradigms be combined when logical frameworks themselves should suffice to formalize and implement logics? We contend, and we hope our development illustrates, that the combination is sensible and advantageous since each paradigm can provide something that the other lacks. On one hand, an LDS can help tailor the consequence relation of a logic to fit better that of the metalogic. On the other, a logical framework provides a means of directly implementing certain kinds of LDS presentations (see discussion in Section 6.2) as ND proof systems, provides a concrete metalogic for reasoning about the correctness of the implementation, and may, as in the case of Isabelle, support structured theory development. Below we consider these points in more detail.

Many of the framework logics which have been actively studied, e.g. the type theory of the Edinburgh LF [10], the higher-order logic of Isabelle [15], and even programming languages like λ-Prolog [6], lend themselves best to representing logics which can be presented as collections of rules for proof under assumption. An example of such a rule is the standard arrow (implication) introduction rule:

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B} \to I$$

This rule is associated with *natural* deduction, which, as the name suggests, is commonly recognized as one of the most natural systems for building proofs, at least for humans (as opposed to computers).

Unfortunately, modal logics fit natural deduction poorly; they are usually presented as Hilbert systems, even though these are recognised as one of the least natural systems for building proofs. This is not to say that it is impossible to give a natural deduction presentation of a modal logic, they have been developed and studied; the problem is that the resulting systems are much more awkward. For instance in any ND presentation of a modal logic based on K, where we have $\to I$, we also are allowed to use the rule

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

where $\Box\Gamma$ indicates that each assumption in Γ has $\Box$ as its outermost connective. The problem with this rule is that it is not *pure*: it carries a side condition on the complete set of assumptions. While logical frameworks work well in encoding certain kinds of rules, namely those rules of ordinary pure single-conclusioned ND systems[1], the logical frameworks so far proposed are not able to formalize the above kind of impure side condition in a natural deduction setting and hence cannot directly formalize such presentations.

The inability to encode impure rules in a logical framework forbids building proof systems using $\to I$ and $\Box I$ together, but not ND presentations of modal logics in general: a pure presentation of $S4$ for the Edinburgh LF logical framework can be found in [2, §4.4], where two judgements (*true* and *valid*) are used which, in essence, factor the proof system into two parts, in one of which only propositional reasoning is possible. While it may be possible to develop other presentations in this fashion, there does not appear to be a systematic way to do this; each new modal logic requires some insight and its own justification of correctness. Further, even when given such presentations, we have no reason to expect them to have the same combinational properties as their corresponding Hilbert systems; i.e. given systems corresponding to $K4$ and KT (i.e. T), we do not know if their combination corresponds to $KT4$ (i.e. $S4$).

We show that the LDS approach can serve as a solution to this problem; for modal logics, it provides precisely what is needed, namely an ordinary, pure single-conclusioned natural deduction presentation. Moreover, the solution supports modularity since the labelling algebra directly expresses the properties of the appropriate Kripke frames.

Finding a 'good' presentation

In order to provide an LDS formalization of a logic we need two things: a base logic, and a general notion of a labelling algebra. However, for each of these there may be more than one possible candidate. For instance in this paper we concentrate on labelling algebras corresponding to Horn theories of the accessibility relation, one possibility out of many, and not even perhaps the most obvious — why restrict ourselves to Horn clause logic, instead of full first-order, or even higher-order, logic?

Clearly we need some criteria for assessing the relative merits of the range of possibilities. We can, of course, consider the basic metatheoretic properties that any logical system is expected to satisfy, such as proof nor-

[1] In [1, footnote to §5.5], Avron summarizes this when he says that "every ordinary, pure single-conclusioned ND system can, e.g., quite easily be implemented on the Edinburgh LF." Note that 'ordinary' means that the system admits the well known rules for contraction and thinning of assumptions.

malization, but we can extend this list. There are pragmatic considerations, such as 'is it easy to use?'. But there are other theoretical considerations: for instance D'Agostino and Gabbay, in [5, p.244], write

> The labelling algebra represents this metalevel information as a *separate* component of a standard derivation system and can be treated as an independent parameter. In the LDS approach, logical systems are not studied statically, in isolation, but dynamically, observing the process of their generation and their interaction (via modifications of the labelling algebras) on the basis of a fixed proof-theoretical hard core (the underlying system of deduction). [their emphasis]

In other words, a good LDS presentation should correspond not just to some logic, but to a space of possible logics, which vary in a well-behaved way according to the details of the labelling algebra; e.g. we would expect that given an LDS for modal logic, a presentation of $K4$ combined with a presentation of T does result in $S4$. By this standard, for instance, while the presentation of $S4$ in [2] could be seen as an LDS where the two judgements correspond to labels, it would not be a good one, since there is no labelling algebra to vary.[2]

The system we propose does well by these measures. It cleanly separates the labelling algebra from the base logic K, and we show that it has good modular, compositional properties for the labelling algebra, behaving in the way we would expect as we combine labelling algebras together, providing a natural hierarchy of systems that inherit theorems and derived rules. Moreover, we use the parameterized relational theory to prove a parameterized completeness theorem with respect to Kripke semantics, and to prove the correctness of the encodings. These theorems show that our implementation not only properly captures modal provability within our hierarchy, but also a satisfactory notion of proof under assumption, i.e. consequence. Third, although not formally quantifiable, our experience shows that proof construction using our presentation is natural and intuitive. Finally, we consider the metatheory of our system, and compare it with other related possibilities, including *semantic embedding*, where the Kripke semantics is used to translate modal propositions into a first-order or higher-order logic.

We show that using our base logic K we are able to interpret the 'separate' in the previous quotation in a strong way: not only do we have a separation between the base logic and the labelling algebra, but that separation is maintained even when building proofs; i.e. the proofs themselves

[2] We do not mean this as a criticism of that presentation, which was not motivated by such concerns.

consist of a derivation tree built from the base logic, which is decorated with a fringe of derivations in the labelling algebra alone. It turns out that this property is directly related to the behavior of *falsum* ($\perp$) in K, which is able to propagate between different worlds. We call this property a *global falsum*. We show that this is enough to implement, among others, the logics in the Geach hierarchy (including many of the modal logics we are likely to encounter in practice), but not enough to implement all modal logics corresponding to first-order definable frames.

Having identified this property of K, we can vary it to produce different candidate 'hard cores'. We investigate the other two obvious possibilities. The first of these, an extension we call *universal falsum*, allows $\perp$ to propagate not only from one world to another, but also between worlds and the labelling algebra (assuming that the labelling algebra is also extended with this). The second, a restriction where $\perp$ is no longer able to propagate even between worlds, we call *local falsum*.

A system with a universal falsum is certainly more general than K. In fact we show that it is equivalent to a traditional semantic embedding in first-order logic, and therefore able to treat not just, e.g., the Geach logics, but any first-order axiomatizable theory. However in exchange for this greater scope we loose the better behaved proof theory of K, and the result does not seem to offer any advantages over semantic embedding in first-order logic (where there is no separation at all), and thus provides no essential alternative to this better known approach. If we restrict ourselves to a local falsum on the other hand, the proof system is in general not suitable for formalizing modal logics, and proofs even no longer have normal forms. Thus K seems to be the weakest logic in which we can embed a useful range of modal logics.

Organization

The remainder of this paper is organized as follows. In Section 2 we present a hierarchy of labelled (propositional) modal logics based on K and Horn relational theories. In Section 3 we consider the soundness and completeness of these theories with respect to Kripke semantics. After, in Section 4, we consider some of the proof-theoretic properties of our encodings and use that to contrast our approach with related formalizations. In Section 5 we sketch our implementation in Isabelle, its application, and its correctness. In Section 6 we compare with related work based on natural deduction presentations of modal logics, LDS presentations, and translation into first-order logic. Finally, we draw conclusions. An Appendix contains proof scripts from an Isabelle session which demonstrate interactive proof construction with our implementation.

2 A Hierarchy of Labelled Modal Logics

We introduce a labelled ND system for the base modal logic K and extend it with (Horn) relational theories.

2.1 The Base Modal Logic K

Definition 2.1 *Let W be a set of* labels *and R a binary relation over W. If x and y are labels, and A is a propositional modal formula built from $\perp$, $\rightarrow$, $\Box$, $\Diamond$, then $x \; R \; y$ is a* relational formula *(*rwff*), and $x : A$ is a* labelled formula *(*lwff*).*

Hence, if p is a sentence letter, and A, B are propositional modal formulae, then $x : p$, $x : \perp$, $x : A \rightarrow B$, $x : \Box A$, $x : \Diamond A$ are all lwffs. Lwffs over other connectives (e.g. $\neg$, $\wedge$, $\vee$) can be defined in the usual manner, e.g. $x : \neg A \equiv x : A \rightarrow \perp$. Henceforth, we assume that the variables x, y, z, w range over labels, the variables A, B range over propositional modal formulae, φ is an arbitrary rwff or lwff, and $\Gamma = \{x_1 : A_1, \ldots, x_n : A_n\}$ and $\Delta = \{x_1 \; R \; y_1, \ldots, x_m \; R \; y_m\}$ are arbitrary sets of lwffs and rwffs. We will also freely use subscripts or superscripts for all of them.

The rules given in Figure 1 determine K, the base ND system which formalizes a labelled version of the modal logic K. For the sake of simplicity, in the following we will sometimes use the rules for negation, $\neg I$ and $\neg E$, which are special cases of $\rightarrow I$ and $\rightarrow E$, respectively:

$$\dfrac{\begin{matrix}[x : A] \\ \vdots \\ x : \perp\end{matrix}}{x : \neg A} \neg I \qquad \dfrac{x : \neg A \quad x : A}{x : \perp} \neg E$$

2.2 Relational Theories

We will formalize particular modal logics by extending K with *relational theories*, which axiomatize properties of the accessibility relation R in Kripke frames. Correspondence theory [21, 22] provides a tool for telling us which modal axioms correspond to which axioms for R. For example, the T axiom,

$\dfrac{\begin{matrix}[x:A\to\bot]\\ \vdots\\ y:\bot\end{matrix}}{x:A}\,\bot E$	$\dfrac{\begin{matrix}[x:A]\\ \vdots\\ x:B\end{matrix}}{x:A\to B}\to I$ $\dfrac{x:A \quad x:A\to B}{x:B}\to E$	$\dfrac{\begin{matrix}[x\;R\;y]\\ \vdots\\ y:A\end{matrix}}{x:\Box A}\,\Box I$ $\dfrac{x:\Box A \quad x\;R\;y}{y:A}\,\Box E$	$\dfrac{y:A \quad x\;R\;y}{x:\Diamond A}\,\Diamond I$ $\dfrac{x:\Diamond A \qquad \begin{matrix}[y:A]\;[x\;R\;y]\\ \vdots\\ z:B\end{matrix}}{z:B}\,\Diamond E$

In $\Box I$ [$\Diamond E$], y is different from x [x and z] and does not occur in the assumptions on which $y:A$ [$z:B$] depends, except those of the form $x\;R\;y$ [$y:A$ and $x\;R\;y$], which are discharged by the inference. We do not enforce Prawitz's side condition on $\bot E$ that $A \neq \bot$.

Figure 1: The rules of K

$\Box A \to A$, corresponds to the first order axiom $\forall x(x\;R\;x)$. Not all modal axioms can be captured in a first-order setting (e.g. the McKinsey axiom $\Box\Diamond A \to \Diamond\Box A$), so there is an important decision that we must make: Should we allow all higher-order relational theories, or some subset thereof?

This decision is non-trivial. We show in Section 4 that different choices of *interface* between K and the labelling algebra result in essentially different systems. Our choice is based on our intention to implement these theories (Section 5.1) as sets of proof rules using a metalogic corresponding to minimal implicational predicate logic. Hence, we have chosen to admit precisely those theories of R which can be directly formulated in the Horn-fragment of this metalogic without requiring additional axioms (e.g. for auxiliary predicates) or judgements (e.g. for identity). We partially justify this choice below by showing that it captures a large class of well-known modal logics including most of those used in practice.

2.3 Horn Relational Theories

Definition 2.2 *A* Horn relational formula *is a closed formula of the form*

$$\forall x_1 \ldots \forall x_n((t_1\;R\;s_1 \wedge \ldots \wedge t_m\;R\;s_m) \to t_0\;R\;s_0)\,,$$

where the t_i and s_i are terms built from the labels $x_1, \ldots, x_n$ and function *symbols. Corresponding to each such formula is a* Horn relational rule

$$\frac{t_1 \; R \; s_1 \quad \ldots \quad t_m \; R \; s_m}{t_0 \; R \; s_0}$$

A Horn relational theory $\mathcal{T}$ *is then a theory generated by a set of such rules.*

In first-order logic the addition of a Horn formula to a theory is equivalent to adding the corresponding rule; hence, in the context of our metatheories we shall talk about additions based on either formulae or rules as is convenient.

We now indicate that restricting our attention to Horn theories is often sufficient in practice. Let i, j, m, and n be natural numbers, and let $\Box^n$ [$\Diamond^n$] stand for a sequence of n consecutive $\Box$s [$\Diamond$s]; for example $\Diamond^2\Box^3\Diamond^0 A$ is $\Diamond\Diamond\Box\Box\Box A$. A large and important class of modal logics falls under the *generalized Geach axiom schema*

$$\Diamond^i \Box^m A \to \Box^j \Diamond^n A \, ,$$

which corresponds to the semantic notion of (i, j, m, n) *convergency* (or 'incestuality' in the terminology of [4])

$$\forall x \forall y \forall z (x \; R^i \; y \wedge x \; R^j \; z \to \exists u (y \; R^m \; u \wedge z \; R^n \; u)) \, ,$$

where $x \; R^0 \; y$ means $x = y$ and $x \; R^{i+1} \; y$ means $\exists v(x \; R \; v \wedge v \; R^i \; y)$.

There are instances of (i, j, m, n) convergency which explicitly require the identity predicate, e.g. $(1, 0, 0, 0)$ yields *vacuity*, $\forall x \forall y (x \; R \; y \to x = y)$. For simplicity, we will not consider theories with identity, and we introduce the subclass of *restricted* (i, j, m, n) *convergency axioms*, as the class of properties of the accessibility relation which can be expressed as Horn rules in the theory of one binary predicate R. These theories yield, among others, most of the modal logics usually of actual interest (K, D, T, B, $S4$, $S4.2$, $KD45$, $S5$,...).

Definition 2.3 Restricted (i, j, m, n) convergency axioms *are closed formulae of the form* $\forall x \forall y \forall z ((x \; R^i \; y \wedge x \; R^j \; z) \to \exists u (y \; R^m \; u \wedge z \; R^n \; u))$*, where* $m = n = 0$ *implies* $i = j = 0$*.*

Proposition 2.4 *If* T_G *is a theory corresponding to a collection of restricted* (i, j, m, n) *convergency axioms, then there is a Horn relational theory* $\mathcal{T}_H$ *conservatively extending it.*

Proof The restriction that $m = n = 0$ implies $i = j = 0$ is a necessary and sufficient condition for identity to be inessential (the necessity can be checked semantically), as noted in [20]. Now, for each convergency axiom A^k in T_G, let B^k be formed by prenexing quantifiers followed by skolemizing remaining existential quantifiers. B^k must be of the form:

$$\forall x_1 \ldots \forall x_l ((t_1\ R\ s_1 \wedge \ldots \wedge t_p\ R\ s_p) \to (t'_1\ R\ s'_1 \wedge \ldots \wedge t'_q\ R\ s'_q)) \,,$$

where $q = m + n \neq 0$, and where Skolem functions only occur in the consequent. We can translate B^k into q Horn relational formulae, B^k_r for $r \in \{1, \ldots, q\}$, of the form

$$\forall x_1 \ldots \forall x_l ((t_1\ R\ s_1 \wedge \ldots \wedge t_p\ R\ s_p) \to t'_r\ R\ s'_r) \,.$$

Let $\mathcal{T}_H$ be the theory generated by the union of the B^k_r rules; the conservativity of $\mathcal{T}_H$ follows by the theorem on functional extensions [19, p.55], and the observation that Skolem constants only occur positively in the B^k_r. (Alternatively, cf. Theorem 3.4.4.(i) in [23, p.137]). □

Some properties corresponding to instances of restricted (i, j, m, n) convergency are given in Figure 2. We also present there the Horn relational rules that result by applying the above translation to these axioms, together with the corresponding characteristic axioms.

Various combinations of Horn relational rules define labelled equivalents of standard propositional modal logics: the logic $L = K + \mathcal{T}$ is obtained by extending K with a given Horn relational theory $\mathcal{T}$.[3] Figure 3 shows a fragment of the resulting hierarchical dependency. For example, $KT4$ ($S4$) is obtained by extending K with the rules R_refl and R_trans, or alternatively by extending either KT with R_trans or $K4$ with R_refl.

Our approach of presenting logics by combinations of K with a relational theory $\mathcal{T}$ provides a general method for representing logics in a modular and transparent way. The relational theory can be viewed as an independent parameter: the base logic K stays fixed for a given class of related logics and we generate the one we want by combining K with the appropriate relational theory. In Section 4, we return to the question of extensions to full first-order or higher-order theories. It is possible to generalize our presentation

[3] We adopt the convention of naming the modal logic $K + \mathcal{T}$ as KAx, where Ax is a string consisting of the standard names of the characteristic axioms corresponding to the relational rules contained in $\mathcal{T}$. As an example, KD, KT, KTB, $KT4$, $KT5$ identify the logics also known as D, T, B, $S4$, $S5$.

Property	(i,j,m,n)	**Char. Axiom**	**Horn Relational Rule**
Seriality	$(0,0,1,1)$	$D: \ \Box A \rightarrow \Diamond A$	$\dfrac{}{x\ R\ f(x)}\ R\_ser$
Reflexivity	$(0,0,1,0)$	$T: \ \Box A \rightarrow A$	$\dfrac{}{x\ R\ x}\ R\_refl$
Symmetry	$(0,1,0,1)$	$B: \ A \rightarrow \Box\Diamond A$	$\dfrac{x\ R\ y}{y\ R\ x}\ R\_symm$
Transitivity	$(0,2,1,0)$	$4: \ \Box A \rightarrow \Box\Box A$	$\dfrac{x\ R\ y \quad y\ R\ z}{x\ R\ z}\ R\_trans$
Euclideaness	$(1,1,0,1)$	$5: \ \Diamond A \rightarrow \Box\Diamond A$	$\dfrac{x\ R\ y \quad x\ R\ z}{z\ R\ y}\ R\_eucl$
Convergency	$(1,1,1,1)$	$2: \ \Diamond\Box A \rightarrow \Box\Diamond A$	$\dfrac{x\ R\ y \quad x\ R\ z}{y\ R\ g(x,y,z)}\ R\_conv1$ $\dfrac{x\ R\ y \quad x\ R\ z}{z\ R\ g(x,y,z)}\ R\_conv2$

Where $f : W \rightarrow W$ and $g : (W \times W \times W) \rightarrow W$ are (Skolem) function constants.

Figure 2: Some properties of R, characteristic axioms, and Horn relational rules

here, but, perhaps surprisingly, for some extensions the 'interface' between K and the relational theory must be changed if completeness for encoded logics (with respect to their intended Kripke semantics) is to be preserved, and the metatheoretic properties of the system change.

2.4 Derivations

We adapt the standard definition of [16] to define derivations of lwffs and rwffs relative to a given relational theory $\mathcal{T}$ used to extend K.

Definition 2.5 *A* derivation *of an lwff or rwff φ from a set of lwffs Γ and a set of rwffs Δ in a logic $L = K + \mathcal{T}$ is a tree formed using the rules in L, ending with φ and depending only on $\Gamma \cup \Delta$. We write $\Gamma, \Delta \vdash_L \varphi$ when φ can be so derived. A derivation of φ in L depending on the empty set, $\vdash_L \varphi$, is a* proof *of φ in L, and we say that φ is an L-*theorem*.

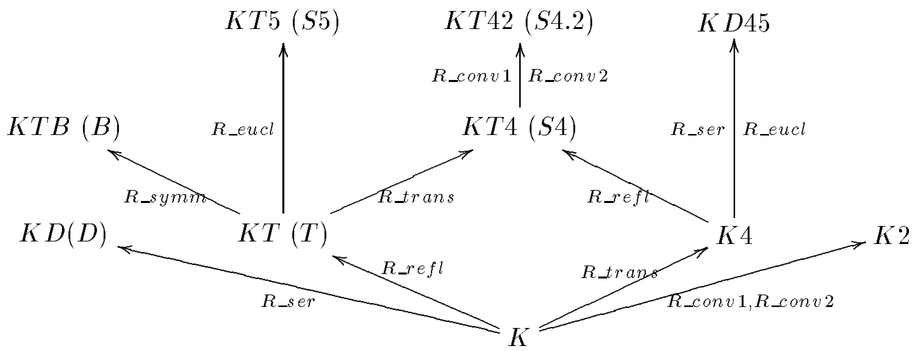

Figure 3: A hierarchy of modal logics (fragment)

Fact 2.6 *When φ is an rwff, say $x\ R\ y$, we have that*

(1) $\Gamma, \Delta \vdash_K x\ R\ y$ *iff* $x\ R\ y \in \Delta$

(2) $\Gamma, \Delta \vdash_{K+\mathcal{T}} x\ R\ y$ *iff* $\Delta \vdash_{K+\mathcal{T}} x\ R\ y$ *iff* $\Delta \vdash_{\mathcal{T}} x\ R\ y$

We also call a derivation [proof] in a logic L an *L-derivation* [*L-proof*], and we will omit the 'L' when the particular logic is not relevant. We systematically use Π, with or without indices, to range over derivations, and we write $\genfrac{}{}{0pt}{}{\Pi}{\varphi}$ to specify that the formula φ is the conclusion of the derivation Π. Similarly, we write $\genfrac{}{}{0pt}{}{\varphi}{\Pi}$ $[\genfrac{}{}{0pt}{}{[\varphi]}{\Pi}]$ to distinguish a possibly empty set of occurrences of the open [discharged] assumption φ in Π. Moreover, we use superscripts to associate discharged assumptions with rule applications.

As an example, we give the $K2$-proof of the characteristic axiom corresponding to convergency, i.e. $\vdash_{K2} x : \Diamond\Box A \to \Box\Diamond A$.

$$
\dfrac{\dfrac{\dfrac{[x : \Diamond\Box A]^3 \qquad \dfrac{\dfrac{[y : \Box A]^1 \qquad \dfrac{[x\ R\ y]^1 \quad [x\ R\ z]^2}{y\ R\ g(x,y,z)}\,R\_conv1}{g(x,y,z) : A}\,\Box E \qquad \dfrac{[x\ R\ y]^1 \quad [x\ R\ z]^2}{z\ R\ g(x,y,z)}\,R\_conv2}{z : \Diamond A}\,\Diamond I}{z : \Diamond A}\,\Diamond E^1}{x : \Box\Diamond A}\,\Box I^2}{x : \Diamond\Box A \to \Box\Diamond A}\,\to I^3
$$

An Isabelle proof for this theorem is presented in Appendix A. As a further example, taken from [8, p.36], we present the K-derivation of $x : \Diamond\Diamond B$ from the assumptions $x : \Box\Box A$, $y : \Diamond(A \to B)$, and $x\ R\ y$.

$$
\begin{prooftree}
\AxiomC{$y\!:\!\Diamond(A \to B)$}
\AxiomC{$x\!:\!\Box\Box A$}
\AxiomC{$x \mathrel{R} y$}
\RightLabel{$\Box E$}
\BinaryInfC{$y\!:\!\Box A$}
\AxiomC{$[y \mathrel{R} z]^1$}
\RightLabel{$\Box E$}
\BinaryInfC{$z\!:\!A$}
\AxiomC{$[z\!:\!A \to B]^1$}
\RightLabel{$\to E$}
\BinaryInfC{$z\!:\!B$}
\AxiomC{$[y \mathrel{R} z]^1$}
\RightLabel{$\Diamond I$}
\BinaryInfC{$y\!:\!\Diamond B$}
\RightLabel{$\Diamond E^1$}
\BinaryInfC{$y\!:\!\Diamond B$}
\AxiomC{$x \mathrel{R} y$}
\RightLabel{$\Diamond I$}
\BinaryInfC{$x\!:\!\Diamond\Diamond B$}
\end{prooftree}
$$

3 Correctness of Labelled Modal Logics

We introduce a Kripke semantics for our systems and modularly prove that any logic L obtained by extending K with a Horn relational theory $\mathcal{T}$ is sound and complete with respect to it.

Definition 3.1 *A* (Kripke) frame *is a pair* (w, r)*, where* w *is a non-empty set, and* $\mathrm{r} \subseteq \mathrm{w} \times \mathrm{w}$*. A* (Kripke) model M *is a triple* $(\mathrm{w}, \mathrm{r}, \mathrm{v})$*, where* (w, r) *is a frame, and* v *maps an element of* w *and a sentence letter to a truth value (0 or 1). A model [frame] is said to have some property of binary relations (e.g. transitivity) iff* r *has that property.*

Note that our models do not contain functions corresponding to possible Skolem functions in the signature. When such constants are present the appropriate Skolem expansion of the model (cf. [23, p.137]) is required.

Definition 3.2 *Given a set of lwffs* Γ *and a set of rwffs* Δ*, we call the ordered pair* (Γ, Δ) *a* proof context (pc)*. When* $\Gamma_1 \subseteq \Gamma_2$ *and* $\Delta_1 \subseteq \Delta_2$*, we write* $(\Gamma_1, \Delta_1) \subseteq (\Gamma_2, \Delta_2)$*, and say that* (Γ_1, Δ_1) is included in (is a subpc of) (Γ_2, Δ_2)*. When* $w\!:\!A \in \Gamma$*, we write* $w\!:\!A \in (\Gamma, \Delta)$ *irrespective of* Δ*, and when* $x \mathrel{R} y \in \Delta$*, we write* $x \mathrel{R} y \in (\Gamma, \Delta)$ *irrespective of* Γ*. Finally, we say that a label* x occurs in (Γ, Δ)*, and by abuse of notation write* $x \in (\Gamma, \Delta)$*, if there exists an* A *such that* $x : A \in \Gamma$*, or a* y *such that* $x \mathrel{R} y \in \Delta$ *or* $y \mathrel{R} x \in \Delta$*.*

Definition 3.3 Truth *for an rwff or lwff* φ *in a model* M*,* $\models^{\mathrm{M}} \varphi$*, is the smallest relation* $\models^{\mathrm{M}}$ *satisfying:*

$$
\begin{array}{lll}
\models^{\mathrm{M}} x \mathrel{R} y & \textit{if} & (x, y) \in \mathrm{r} \\
\models^{\mathrm{M}} x\!:\!p & \textit{if} & \mathrm{v}(x, p) = 1 \\
\models^{\mathrm{M}} x\!:\!A \to B & \textit{if} & \models^{\mathrm{M}} x\!:\!A \textit{ implies} \models^{\mathrm{M}} x\!:\!B \\
\models^{\mathrm{M}} x\!:\!\Box A & \textit{if} & \textit{for all } y, \models^{\mathrm{M}} x \mathrel{R} y \textit{ implies} \models^{\mathrm{M}} y\!:\!A \\
\models^{\mathrm{M}} x\!:\!\Diamond A & \textit{if} & \textit{for some } y, \models^{\mathrm{M}} x \mathrel{R} y \textit{ and} \models^{\mathrm{M}} y\!:\!A
\end{array}
$$

When $\models^{\mathrm{M}} \varphi$*, we say that* φ *is* true *in* M*. By extension,* $\models^{\mathrm{M}} (\Gamma, \Delta)$ *means that* $\models^{\mathrm{M}} \varphi$ *for all* $\varphi \in (\Gamma, \Delta)$*, and* $\Gamma, \Delta \models \varphi$ *means that* $\models^{\mathrm{M}} (\Gamma, \Delta)$ *implies* $\models^{\mathrm{M}} \varphi$ *for any model* M*.*

Truth for lwffs containing other connectives, e.g. $\models^{\mathrm{M}} x : \neg A$, can be defined in the usual manner. Moreover, truth for lwffs is related to the standard truth relation for unlabelled modal logics, e.g. [4], by observing that $\models^{\mathrm{M}} x : A$ iff $\models^{\mathrm{M}}_x A$. Analogous to Fact 2.6 we have that:

Fact 3.4 $\Gamma, \Delta \models x\ R\ y$ *iff* $\Delta \models x\ R\ y$*.*

Definition 3.5 *The modal logic* $L = K + \mathcal{T}$ *is* sound *iff* $\Gamma, \Delta \vdash_L \varphi$ *implies* $\Gamma, \Delta \models \varphi$*.* L *is* complete *iff the converse holds.*

The explicit embedding of properties of the models, and the possibility of explicitly reasoning about them, via rwffs and relational rules, require us to consider also soundness and completeness for rwffs, where, by Fact 2.6 and Fact 3.4, we show that $\Delta \vdash_L x\ R\ y$ iff $\Delta \models x\ R\ y$.

Lemma 3.6 $L = K + \mathcal{T}$ *is sound, i.e.*

(1) $\Delta \vdash_L x\ R\ y$ *implies* $\Delta \models x\ R\ y$

(2) $\Gamma, \Delta \vdash_L x : A$ *implies* $\Gamma, \Delta \models x : A$

Proof Throughout the proof let $\mathrm{M}_L = (\mathrm{w}_L, \mathrm{r}_L, \mathrm{v}_L)$ be an arbitrary model for the logic L. We prove (1) by induction on the structure of the derivation of $x\ R\ y$ from Δ. The base case ($x\ R\ y \in \Delta$) is trivial. There is one step for each Horn relational rule; we treat only transitivity and convergency as examples. For transitivity, assume that r_L is transitive and consider applications of the rule *R_trans*

$$\dfrac{\begin{matrix}\Pi_1 \\ x\ R\ y\end{matrix} \qquad \begin{matrix}\Pi_2 \\ y\ R\ z\end{matrix}}{x\ R\ z}\ R\_trans$$

where Π_1 and Π_2 are the derivations $\Delta_1 \vdash_L x\ R\ y$ and $\Delta_2 \vdash_L y\ R\ z$, with $\Delta = \Delta_1 \cup \Delta_2$. The induction hypotheses are $\Delta_1 \vdash_L x\ R\ y$ implies $\Delta_1 \models x\ R\ y$, and $\Delta_2 \vdash_L y\ R\ z$ implies $\Delta_2 \models y\ R\ z$. Assume $\models^{\mathrm{M}_L} \Delta$. Then, from the induction hypotheses we obtain $\models^{\mathrm{M}_L} x\ R\ y$ and $\models^{\mathrm{M}_L} y\ R\ z$, i.e. $(x, y) \in \mathrm{r}_L$ and $(y, z) \in \mathrm{r}_L$. Since r_L is transitive, we conclude $\models^{\mathrm{M}_L} x\ R\ z$ by Definition 3.3.

In the case of Skolem constants M_L is a Skolem expansion; e.g. consider applications of the rules R_conv1 and R_conv2

$$\dfrac{\begin{array}{cc}\Pi_1 & \Pi_2 \\ x\ R\ y & x\ R\ z\end{array}}{y\ R\ g(x,y,z)}\ R\_conv1 \qquad \dfrac{\begin{array}{cc}\Pi_1 & \Pi_2 \\ x\ R\ y & x\ R\ z\end{array}}{z\ R\ g(x,y,z)}\ R\_conv2$$

where Π_1 and Π_2 are the derivations $\Delta_1 \vdash_L x\ R\ y$ and $\Delta_2 \vdash_L x\ R\ z$, with $\Delta = \Delta_1 \cup \Delta_2$. By Proposition 2.4, the theory $\mathcal{T}_H$ generated by R_conv1 and R_conv2 is a conservative extension of the first-order theory T_G corresponding to the convergency axiom. By Theorem 3.4.4.(ii) in [23, p.137], each model of the theory T_G has a Skolem expansion, contained in M_L, which is a model of $\mathcal{T}_H$. Assume $\models^{\mathrm{M}_L} \Delta$. Then, from the induction hypotheses we obtain $\models^{\mathrm{M}_L} x\ R\ y$ and $\models^{\mathrm{M}_L} x\ R\ z$, i.e. $(x,y) \in \mathrm{r}_L$ and $(x,z) \in \mathrm{r}_L$. Since r_L is convergent, we have $\models^{\mathrm{M}_L} y\ R\ g(x,y,z)$ and $\models^{\mathrm{M}_L} z\ R\ g(x,y,z)$ by Definition 3.3.

We prove (2) by induction on the structure of the derivation of $x : A$ from Γ and Δ. The base case ($x : A \in \Gamma$) is trivial. There is one step for each inference rule; we treat only applications of $\bot E$, $\Box I$ and $\Box E$.

$\boxed{\bot E}$

$$\dfrac{\begin{array}{c}[x : A \to \bot] \\ \Pi \\ y : \bot\end{array}}{x : A}\ \bot E$$

where Π is the derivation $\Gamma_1, \Delta \vdash_L y : \bot$, with $\Gamma_1 = \Gamma \cup \{x : A \to \bot\}$. The induction hypothesis is $\Gamma_1, \Delta \vdash_L y : \bot$ implies $\Gamma_1, \Delta \models y : \bot$. We assume $\models^{\mathrm{M}_L} (\Gamma, \Delta)$, and prove $\models^{\mathrm{M}_L} x : A$. Since $\not\models^{\mathrm{M}_L} y : \bot$ for any y, from the induction hypothesis we obtain $\not\models^{\mathrm{M}_L} \Gamma_1$, and therefore $\not\models^{\mathrm{M}_L} \{x : A \to \bot\}$, i.e. $\models^{\mathrm{M}_L} x : A$ and $\not\models^{\mathrm{M}_L} x : \bot$ by Definition 3.3.

$\boxed{\Box I}$

$$\dfrac{\begin{array}{c}[x\ R\ y] \\ \Pi \\ y : A\end{array}}{x : \Box A}\ \Box I$$

where Π is the derivation $\Gamma, \Delta_1 \vdash_L y : A$, with $\Delta_1 = \Delta \cup \{x\ R\ y\}$. The induction hypothesis is $\Gamma, \Delta_1 \vdash_L y : A$ implies $\Gamma, \Delta_1 \models y : A$. Assume

$\models^{\mathrm{M}_L} (\Gamma, \Delta)$. Considering the restriction on the application of $\Box I$, we can extend Δ to $\Delta' = \Delta \cup \{x\ R\ z\}$ for an arbitrary $z \notin (\Gamma, \Delta)$, and assume $\models^{\mathrm{M}_L} \Delta'$. Since $\models^{\mathrm{M}_L} \Delta'$ implies $\models^{\mathrm{M}_L} \Delta_1$, from the induction hypothesis we obtain $\models^{\mathrm{M}_L} y : A$, that is $\models^{\mathrm{M}_L} z : A$ for an arbitrary $z \notin (\Gamma, \Delta)$ such that $\models^{\mathrm{M}_L} x\ R\ z$. We conclude $\models^{\mathrm{M}_L} x : \Box A$ by Definition 3.3.

$\boxed{\Box E}$

$$\frac{\begin{array}{c}\Pi_1 \\ x : \Box A\end{array} \qquad \begin{array}{c}\Pi_2 \\ x\ R\ y\end{array}}{y : A}\ \Box E$$

where Π_1 and Π_2 are the derivations $\Gamma, \Delta_1 \vdash_L x : \Box A$ and $\Delta_2 \vdash_L x\ R\ y$, with $\Delta = \Delta_1 \cup \Delta_2$. Assume $\models^{\mathrm{M}_L} (\Gamma, \Delta)$. Then, from the induction hypotheses we obtain $\models^{\mathrm{M}_L} x : \Box A$ and $\models^{\mathrm{M}_L} x\ R\ y$, and thus $\models^{\mathrm{M}_L} y : A$ by Definition 3.3. $\Box$

Definition 3.7 *Let $L = K + \mathcal{T}$ be a consistent logic, i.e. $\not\vdash_L x : \bot$ for every label x. A pc (Γ, Δ) is L-*consistent *iff $\Gamma, \Delta \not\vdash_L x : \bot$ for every label x. (Γ, Δ) is L-*inconsistent *iff it is not L-consistent.*

When the particular logic is not relevant, we will omit the 'L' and simply speak of consistent and inconsistent pcs.

Fact 3.8 *If (Γ, Δ) is consistent, then for every lwff $x : A$ either $(\Gamma \cup \{x : A\}, \Delta)$ is consistent or $(\Gamma \cup \{x : \neg A\}, \Delta)$ is consistent.*

For any logic $L = K + \mathcal{T}$, let Δ_L be the *deductive closure* of Δ under $\mathcal{T}$, i.e.

$$\Delta_L = \{x\ R\ y \mid \Delta \vdash_L x\ R\ y\}\,.$$

Note that $\Gamma, \Delta \vdash_L \varphi$ iff $\Gamma, \Delta_L \vdash_L \varphi$, and that Δ_L might be empty when Δ is empty.

Definition 3.9 *A pc (Γ, Δ) is* maximally consistent *iff (1) it is consistent; (2) $\Delta = \Delta_L$; (3) for every $x : A$ either $x : A \in (\Gamma, \Delta)$ or $x : \neg A \in (\Gamma, \Delta)$.*

Completeness follows by a modification of the standard Henkin-style proof, where a canonical model $\mathrm{M}_L^C = (\mathrm{w}_L^C, \mathrm{r}_L^C, \mathrm{v}_L^C)$ is built to show that[4]

[4] We consider only consistent pcs. If (Γ, Δ) is inconsistent, then $\Gamma, \Delta \vdash_L x : A$ for all $x : A$, and thus completeness immediately holds for lwffs. Our labelling algebra does not allow us to define inconsistency for a set of rwffs, but, if (Γ, Δ) is inconsistent, the canonical model built in the following is nonetheless a countermodel to non-derivable rwffs.

$$\Gamma, \Delta \not\vdash_L \varphi \text{ implies } \Gamma, \Delta \not\models^{\mathrm{M}_L^C} \varphi.$$

In standard proofs for unlabelled modal logics the set w_L^C is obtained by progressively building maximally consistent sets of formulae, where consistency is locally checked within each set (cf. [4]). In our case, given the presence of labelled formulae and explicit assumptions on the relations between the labels, i.e. Δ, we modify the Lindenbaum lemma (Lemma 3.10 below) to extend (Γ, Δ) to one single maximally consistent proof context (Γ^*, Δ^*), where consistency is 'globally' checked also against the additional assumptions in Δ. The elements of w_L^C are then built by partitioning Γ^* with respect to the labels, and accessibility is defined by exploiting the information in Δ^*. Moreover, in standard proofs the way in which w_L^C is built depends on the particular modal logic L, in particular on the accessibility conditions holding for L. In our case, the proof is completely independent of L: exactly the same procedure applies for any logic.

In the Lindenbaum lemma for first-order logic a maximally consistent and ω-complete set of formulae is inductively built by adding for every formula $\exists x.P(x)$ a *witness* to its truth, namely a formula $P(c)$ for some new individual constant c. This ensures that if, for every closed term t, $P(t)$ is contained in the set, then so is $\forall x.P(x)$. A similar procedure applies here in the case of lwffs of the form $x : \Diamond A$. That is, together with $x : \Diamond A$ we consistently add $y : A$ and $x \; R \; y$ for some new y, which acts as a witness world to the truth of $x : \Diamond A$. This ensures that the maximally consistent pc (Γ^*, Δ^*) is such that if $x \; R \; z \in (\Gamma^*, \Delta^*)$ implies $z : B \in (\Gamma^*, \Delta^*)$ for every z, then $x : \Box B \in (\Gamma^*, \Delta^*)$, as shown in Lemma 3.11 below. Note that in the standard completeness proof for unlabelled modal logics, one shows instead that for every $w \in \mathrm{w}_L^C$, if $\Diamond A \in w$, then w_L^C also contains a world accessible from w that serves as a witness world to the truth of $\Diamond A$.

Lemma 3.10 *Every consistent pc* (Γ, Δ) *can be extended to a maximally consistent pc* (Γ^*, Δ^*).

Proof We first extend the language of the logic L with infinitely many new constants for witness worlds. Systematically let w range over labels, v range over the new constants for witness worlds, and u range over both. All these may be subscripted. Let $l_1, l_2, \ldots$ be an enumeration of all lwffs in the extended language. Starting from $(\Gamma_0, \Delta_0) = (\Gamma, \Delta)$, we inductively build a sequence of consistent pcs by defining $(\Gamma_{i+1}, \Delta_{i+1})$ to be:

- (Γ_i, Δ_i), if $(\Gamma_i \cup \{l_{i+1}\}, \Delta_i)$ is inconsistent; else

- $(\Gamma_i \cup \{l_{i+1}\}, \Delta_i)$, if l_{i+1} is not $u\!:\!\Diamond A$; else
- $(\Gamma_i \cup \{u\!:\!\Diamond A, v\!:\!A\}, \Delta_i \cup \{u\ R\ v\})$, for a $v \notin (\Gamma_i \cup \{u\!:\!\Diamond A\}, \Delta_i)$, if l_{i+1} is $u\!:\!\Diamond A$.

Every (Γ_i, Δ_i) is consistent. We show that if $(\Gamma_i \cup \{u\!:\!\Diamond A\}, \Delta_i)$ is consistent, then so is $(\Gamma_i \cup \{u\!:\!\Diamond A, v\!:\!A\}, \Delta_i \cup \{u\ R\ v\})$, for a $v \notin (\Gamma_i \cup \{u\!:\!\Diamond A\}, \Delta_i)$; the other cases follow by construction. Suppose that for any $v \notin (\Gamma_i \cup \{u : \Diamond A\}, \Delta_i)$, $\Gamma_i \cup \{u : \Diamond A, v : A\}, \Delta_i \cup \{u\ R\ v\} \vdash_L u_j : \bot$. Then $\Gamma_i \cup \{u : \Diamond A\}, \Delta_i \cup \{u\ R\ v\} \vdash_L v : \neg A$, and $\Box I$ yields $\Gamma_i \cup \{u\!:\!\Diamond A\}, \Delta_i \vdash_L u : \Box\neg A$, i.e. $\Gamma_i \cup \{u\!:\!\Diamond A\}, \Delta_i \vdash_L u\!:\!\neg\Diamond A$. Thus $\Gamma_i, \Delta_i \vdash_L u\!:\!\bot$. Contradiction.

Now let $(\Gamma^*, \Delta^*) = (\bigcup_{i\geq 0} \Gamma_i, (\bigcup_{i\geq 0} \Delta_i)_L)$. We show that (Γ^*, Δ^*) is maximally consistent by proving that it satisfies the conditions in Definition 3.9. For (1), note that if $(\bigcup_{i\geq 0} \Gamma_i, \bigcup_{i\geq 0} \Delta_i)$ is consistent, so is $(\bigcup_{i\geq 0} \Gamma_i, (\bigcup_{i\geq 0} \Delta_i)_L)$. Now suppose that (Γ^*, Δ^*) is inconsistent. Then for some finite subpc (Γ', Δ') there exists a u such that $\Gamma', \Delta' \vdash_L u : \bot$. Every lwff $l \in (\Gamma', \Delta')$ is in some (Γ_j, Δ_j). For each $l \in (\Gamma', \Delta')$, let i_l be the least j such that $l \in (\Gamma_j, \Delta_j)$, and let $i = \max\{i_l \mid l \in (\Gamma', \Delta')\}$. Then $(\Gamma', \Delta') \subseteq (\Gamma_i, \Delta_i)$, and (Γ_i, Δ_i) is inconsistent, which is not the case. (2) is satisfied by definition of Δ^*. For (3), suppose that $l_{i+1} \notin (\Gamma^*, \Delta^*)$. Then $l_{i+1} \notin (\Gamma_{i+1}, \Delta_{i+1})$ and $(\Gamma_i \cup \{l_{i+1}\}, \Delta_i)$ is inconsistent. Thus, by Fact 3.8, $(\Gamma_i \cup \{\neg l_{i+1}\}, \Delta_i)$ is consistent, and $\neg l_{i+1}$ is consistently added to some (Γ_j, Δ_j) during the construction, and therefore $\neg l_{i+1} \in (\Gamma^*, \Delta^*)$. □

Lemma 3.11 *Let* (Γ^*, Δ^*) *be a maximally consistent pc. Then*

(1) $\Gamma^*, \Delta^* \vdash_L u_i\ R\ u_j$ *iff* $u_i\ R\ u_j \in (\Gamma^*, \Delta^*)$

(2) $\Gamma^*, \Delta^* \vdash_L u\!:\!A$ *iff* $u\!:\!A \in (\Gamma^*, \Delta^*)$ (deductive closure)

(3) $u\!:\!B \to C \in (\Gamma^*, \Delta^*)$ *iff* $u\!:\!B \in (\Gamma^*, \Delta^*)$ *implies* $u\!:\!C \in (\Gamma^*, \Delta^*)$

(4) $u_i : \Box B \in (\Gamma^*, \Delta^*)$ *iff for all* u_j, $u_i\ R\ u_j \in (\Gamma^*, \Delta^*)$ *implies* $u_j : B \in (\Gamma^*, \Delta^*)$

(5) $u_i : \Diamond B \in (\Gamma^*, \Delta^*)$ *iff for some* u_j, $u_i\ R\ u_j \in (\Gamma^*, \Delta^*)$ *and* $u_j : B \in (\Gamma^*, \Delta^*)$

Proof We only treat (4). Suppose that $u_i : \Box B \in (\Gamma^*, \Delta^*)$. Then, by deductive closure, $\Gamma^*, \Delta^* \vdash_L u_i\!:\!\Box B$, and, by $\Box E$, $\Gamma^*, \Delta^* \vdash_L u_i\ R\ u_j$ implies $\Gamma^*, \Delta^* \vdash_L u_j\!:\!B$ for all u_j. By deductive closure, conclude $u_i\ R\ u_j \in (\Gamma^*, \Delta^*)$ implies $u_j\!:\!B \in (\Gamma^*, \Delta^*)$ for all u_j. For the converse, suppose that $u_i\!:\!\Box B \notin (\Gamma^*, \Delta^*)$. Then $u_i\!:\!\neg\Box B \in (\Gamma^*, \Delta^*)$, i.e. $u_i\!:\!\Diamond\neg B \in (\Gamma^*, \Delta^*)$. Hence, by the construction of (Γ^*, Δ^*), there exists a u_j such that $u_i \; R \; u_j \in (\Gamma^*, \Delta^*)$ and $u_j : B \notin (\Gamma^*, \Delta^*)$. $\Box$

Definition 3.12 *Given* (Γ^*, Δ^*)*, we define the* canonical model M_L^C *for the logic* L *as follows:* $\mathrm{w}_L^C = \{\{A \mid u : A \in \Gamma^*\} \mid u \in (\Gamma^*, \Delta^*)\}$*;* $(u_i, u_j) \in \mathrm{r}_L^C$ *iff* $u_i \; R \; u_j \in \Delta^*$*;* $\mathrm{v}_L^C(u, p) = 1$ *iff* $u : p \in \Gamma^*$*.*

The standard definition of r_L^C, i.e. $(u_i, u_j) \in \mathrm{r}_L^C$ iff $\{A \mid \Box A \in u_i\} \subseteq u_j$, is not applicable in our setting, since $\{A \mid \Box A \in u_i\} \subseteq u_j$ does *not* imply $\vdash_L u_i \; R \; u_j$. We would therefore be unable to prove completeness for rwffs, since there would be cases, e.g. when $L = K$ and $\Delta = \{\}$, where $\not\vdash_L u_i \; R \; u_j$ but $(u_i, u_j) \in \mathrm{r}_L^C$, and thus $\models^{\mathrm{M}_L^C} u_i \; R \; u_j$. Hence, we instead define $(u_i, u_j) \in \mathrm{r}_L^C$ iff $u_i \; R \; u_j \in \Delta^*$; note that therefore $u_i \; R \; u_j \in \Delta^*$ implies $\{A \mid \Box A \in u_i\} \subseteq u_j$. Moreover, we immediately have that:

Fact 3.13 $u_i \; R \; u_j \in \Delta^*$ *iff* $\Delta^* \models^{\mathrm{M}_L^C} u_i \; R \; u_j$*.*

The deductive closure of Δ^* ensures not only completeness for rwffs (as shown in Lemma 3.16 below), but also that the conditions on r_L^C are satisfied, so that M_L^C is really a model for L. As an example, we show that if L contains *R_conv*1 and *R_conv*2, then r_L^C is convergent. Consider an arbitrary pc (Γ, Δ), from which we build M_L^C. Assume $(u_i, u_j) \in \mathrm{r}_L^C$ and $(u_i, u_k) \in \mathrm{r}_L^C$. Then $u_i \; R \; u_j \in \Delta^*$ and $u_i \; R \; u_k \in \Delta^*$. But Δ^* is deductively closed, and thus $u_j \; R \; g(u_i, u_j, u_k) \in \Delta^*$ and $u_k \; R \; g(u_i, u_j, u_k) \in \Delta^*$. Hence, there exists a u_l such that $(u_j, u_l) \in \mathrm{r}_L^C$ and $(u_k, u_l) \in \mathrm{r}_L^C$.

Definition 3.14 *The* degree *of an lwff is the number of times* $\bot$*,* $\to$ *and* $\Box$ *occur in it.*

Lemma 3.15 $u : A \in (\Gamma^*, \Delta^*)$ *iff* $\Gamma^*, \Delta^* \models^{\mathrm{M}_L^C} u : A$*.*

Proof By induction on the degree of $u : A$; we treat only the step case given by $u_i : \Box B$. Assume $u_i : \Box B \in (\Gamma^*, \Delta^*)$. Then, by Lemma 3.11, $u_i \; R \; u_j \in (\Gamma^*, \Delta^*)$ implies $u_j : B \in (\Gamma^*, \Delta^*)$, for all u_j. Fact 3.13 and the induction hypothesis yield $\Gamma^*, \Delta^* \models^{\mathrm{M}_L^C} u_j : B$ for all u_j such that $\Gamma^*, \Delta^* \models^{\mathrm{M}_L^C} u_i \; R \; u_j$, i.e. $\Gamma^*, \Delta^* \models^{\mathrm{M}_L^C} u_i : \Box B$ by Definition 3.3. For the converse, assume $u_i : \neg\Box B \in (\Gamma^*, \Delta^*)$. Then, by Lemma 3.11, $u_i \; R \; u_j \in (\Gamma^*, \Delta^*)$ and $u_j : \neg B \in (\Gamma^*, \Delta^*)$, for some u_j. Fact 3.13 and the induction hypothesis yield $\Gamma^*, \Delta^* \models^{\mathrm{M}_L^C} u_i \; R \; u_j$ and $\Gamma^*, \Delta^* \models^{\mathrm{M}_L^C} u_j : \neg B$, i.e. $\Gamma^*, \Delta^* \models^{\mathrm{M}_L^C} u_i : \neg\Box B$ by Definition 3.3. $\Box$

We can now finally show that:

Lemma 3.16 *$L = K + \mathcal{T}$ is complete, i.e.*

(1) $\Delta \models w_i\ R\ w_j$ *implies* $\Delta \vdash_L w_i\ R\ w_j$

(2) $\Gamma, \Delta \models w\!:\!A$ *implies* $\Gamma, \Delta \vdash_L w\!:\!A$

Proof (1) If $\Delta \not\vdash_L w_i\ R\ w_j$, then $w_i\ R\ w_j \notin \Delta^*$, and thus $\Delta^* \not\models^{\mathrm{M}_L^C} w_i\ R\ w_j$, by Fact 3.13. (2) If $\Gamma, \Delta \not\vdash_L w : A$, then $(\Gamma \cup \{w : \neg A\}, \Delta)$ is consistent. Otherwise there exists a w_i such that $\Gamma \cup \{w\!:\!\neg A\}, \Delta \vdash_L w_i : \perp$, and then $\Gamma, \Delta \vdash_L w : A$. Therefore, by Lemma 3.10, $(\Gamma \cup \{w : \neg A\}, \Delta)$ is included in a maximally consistent pc $((\Gamma \cup \{w\!:\!\neg A\})^*, \Delta^*)$. Then, by Lemma 3.15, $(\Gamma \cup \{w\!:\!\neg A\})^*, \Delta^* \models^{\mathrm{M}_L^C} w\!:\!\neg A$, i.e. $(\Gamma \cup \{w\!:\!\neg A\})^*, \Delta^* \not\models^{\mathrm{M}_L^C} w\!:\!A$, and thus $\Gamma, \Delta \not\models^{\mathrm{M}_L^C} w\!:\!A$. $\Box$

By Lemma 3.6 and Lemma 3.16 we immediately have that:

Theorem 3.17 *$L = K + \mathcal{T}$ is sound and complete.*

4 A Topography of Labelled Modal Logics

We have given a particular presentation of i (propositional) modal logics as Labelled Deductive Systems based on two separate parts: a base logic, K, and Horn relational theories. Here we consider some alternatives for defining hierarchies of logics and classify them based on their metatheoretic properties. We organize this investigation around the interface between the two parts: since the rules for $\Box$ and $\Diamond$ cannot be sensibly changed, this amounts to studying how *falsum* ($\perp$) propogates between worlds. We show that this question directly relates to which kinds of relational theories we can formalize while retaining completeness.

We start in Section 4.1 with the base logic K we have developed above, where we have what we call global falsum: $\perp$ can propagate from one world to another (Fact 4.1). We prove that this system preserves duality between $\Box$ and $\Diamond$ (Proposition 4.2) and that derivations have good normalization properties (Theorem 4.6) in comparison with what we get from semantic embedding (Fact 4.10 and Fact 4.12). These good properties, however, mean that using K we are not able to formalize all modal logics with first-order axiomatizable frames (Theorem 4.11).

In Section 4.2 we consider what happens if we allow $\perp$ to propagate between base logic and labelling algebra in either direction. By doing this, we loose the good normalization properties of K (Fact 4.12) in exchange for a system (K^{uf}, K with universal falsum) that is essentially equivalent to semantic embedding in first-order logic (Theorem 4.14).

Finally, in Section 4.3 we investigate the properties of K^{lf} (K with local falsum), the base logic we get by restricting $\bot E$ in K so that all references are local to one world. Here, unlike in K, we cannot propagate $\bot$ freely from one world to another (Proposition 4.16). We argue that though certain modal logics can be formalized in extensions of K^{lf}, the system lacks basic properties, such as duality between $\Box$ and $\Diamond$ (Proposition 4.18) or normal form derivations (Theorem 4.20), which we might look for in a 'good' formalization.

4.1 Global Falsum

We begin by observing that in K, and therefore in $K + \mathcal{T}$, $\bot$ propagates 'globally' between all worlds. We call this property *global falsum*, and as an immediate consequence of $\bot E$ (where no assumptions are discharged) we have:

Fact 4.1 *The rule* $\dfrac{x:\bot}{y:\bot}\,gf$ *is derivable in* K.

Where possible, we follow Prawitz [16]; like him, we introduce some restrictions to simplify the development. We consider the (functionally complete) $\bot, \rightarrow, \Box$ fragment of the system given in Section 2.1, where we restrict applications of $\bot E$ to the case where the consequence $x:A$ is atomic (i.e. A is atomic). These restrictions are justified by the two following propositions.

Proposition 4.2 *The connectives* $\Box$ *and* $\Diamond$ *are interdefinable in* K.

Proof We define $\Diamond A$ as $\neg\Box\neg A$, and show that the rules for $\Diamond$ are derivable.

$$\dfrac{y:A \quad x\,R\,y}{x:\Diamond A}\,\Diamond I \quad \rightsquigarrow \quad \dfrac{\dfrac{\dfrac{y:A \quad \dfrac{[x:\Box\neg A]^1 \quad x\,R\,y}{y:\neg A}\,\Box E}{y:\bot}\,\neg E}{x:\bot}\,gf}{x:\neg\Box\neg A}\,\neg I^1 \tag{1}$$

$$
\frac{x:\Diamond A \qquad \begin{matrix}[y:A]\ [x\ R\ y]\\ \Pi \\ z:B\end{matrix}}{z:B}\,\Diamond E
\quad\rightsquigarrow\quad
\cfrac{\cfrac{x:\neg\Box\neg A \qquad \cfrac{\cfrac{\cfrac{\cfrac{\begin{matrix}[y:A]^1\ [x\ R\ y]^2\\ \Pi \\ z:B\end{matrix} \qquad [z:B\to\bot]^3}{z:\bot}\,\to E}{y:\bot}\,gf}{y:\neg A}\,\neg I^1}{x:\Box\neg A}\,\Box I^2}{x:\bot}\,\neg E}{z:B}\,\bot E^3
\tag{2}
$$

Dually, we can take $\Diamond$ as primitive and derive the rules for $\Box$. $\Box$

Proposition 4.3 *If $\Gamma, \Delta \vdash_K x:A$, then there is a derivation of $x:A$ from Γ, Δ in the $\bot, \to, \Box$ fragment of K, where the consequences of applications of $\bot E$ are atomic.*

Proof Substitute applications of $\Diamond I$ and $\Diamond E$ as in (1) and (2). We show that any application of $\bot E$ with a non-atomic consequence can be replaced with a derivation in which $\bot E$ is applied only to lwffs of smaller degree. By Proposition 4.2, there are two possible cases, depending on whether the conclusion is $x:A\to B$ or $x:\Box A$.

Case one:

$$
\cfrac{\begin{matrix}[x:(A\to B)\to\bot]\\ \Pi \\ y:\bot\end{matrix}}{x:A\to B}\,\bot E
\quad\rightsquigarrow\quad
\cfrac{\cfrac{\begin{matrix}\cfrac{\cfrac{\cfrac{[x:A]^3 \quad [x:A\to B]^1}{x:B}\,\to E \qquad [x:B\to\bot]^2}{x:\bot}\,\to E}{x:(A\to B)\to\bot}\,\to I^1\\ \Pi \\ y:\bot\end{matrix}}{x:B}\,\bot E^2}{x:A\to B}\,\to I^3
$$

Case two:

$$
\cfrac{\begin{matrix}[x:\Box A\to\bot]\\ \Pi \\ y:\bot\end{matrix}}{x:\Box A}\,\bot E
\quad\rightsquigarrow\quad
\cfrac{\cfrac{\begin{matrix}\cfrac{\cfrac{\cfrac{\cfrac{[x:\Box A]^1 \quad [x\ R\ y]^3}{y:A}\,\Box E \qquad [y:A\to\bot]^2}{y:\bot}\,\to E}{x:\bot}\,gf}{x:\Box A\to\bot}\,\to I^1\\ \Pi \\ y:\bot\end{matrix}}{y:A}\,\bot E^2}{x:\Box A}\,\Box I^3
$$

Conclude by successively repeating the transformation. $\square$

An immediate consequence of this is the equivalence of the restricted and the unrestricted ND system. We will therefore refer to both of them as K.

Definition 4.4 *Any lwff $x:A$ in a derivation is the root of a tree of rule applications leading back to assumptions. The lwffs in this tree other than $x:A$ we call* side lwffs *of $x:A$ in the derivation. A* maximal lwff *in a derivation is an lwff which is both the conclusion of an introduction rule and the major premise of an elimination rule. A maximal lwff can be removed from a derivation by a* reduction step. *Two possible configurations (for $\to$ and $\square$) result in a maximal lwff in a derivation. They, and their corresponding reduction steps are:*

$$\dfrac{\dfrac{\begin{array}{c}[x:A]^1\\ \Pi_1\\ x:B\end{array}}{x:A\to B}\to I^1 \qquad \begin{array}{c}\Pi_2\\ x:A\end{array}}{x:B}\to E \qquad \rightsquigarrow \qquad \begin{array}{c}\Pi_2\\ x:A\\ \Pi_1\\ x:B\end{array} \tag{3}$$

$$\dfrac{\dfrac{\begin{array}{c}[x\ R\ y]^1\\ \Pi\\ y:A\end{array}}{x:\square A}\square I^1 \qquad x\ R\ z}{z:A}\square E \qquad \rightsquigarrow \qquad \begin{array}{c}x\ R\ z\\ \Pi[z/y]\\ z:A\end{array} \tag{4}$$

where $\Pi[z/y]$ is obtained from Π by systematically substituting z for y, with a suitable renaming of the variables to avoid clashes. Note that we only show the part of the derivation where the reduction actually takes place; the missing parts remain unchanged.

Definition 4.5 *A derivation is in* normal form *(is a* normal derivation*) if it contains no maximal lwffs.*

Theorem 4.6 *Every derivation of $x:A$ from Γ,Δ in K reduces to normal form.*

Proof If Π is a derivation of $x:A$ from Γ,Δ in K, then from the set of maximal lwffs of Π pick some $y:B$ which has the highest degree and has maximal lwffs only of lower degree as side lwffs. Let Π' be the reduction of Π at $y:B$. Π' is also a derivation of $x:A$ from Γ,Δ in K and no new maximal lwff as large, or larger than $y:B$ has been introduced. Hence, by a finite number of similar reductions we obtain a derivation of $x : A$ from Γ, Δ in K containing no maximal lwffs. $\Box$

Since derivations in a Horn relational theory $\mathcal{T}_H$ cannot introduce maximal lwffs (and all the rwffs are of the form $x\ R\ y$), by minor modifications to the above, e.g. substitute $\genfrac{}{}{0pt}{}{\Pi_2}{x R z}$ for $x\ R\ z$ in (4), we immediately have:

Corollary 4.7 *Every derivation in $K + \mathcal{T}_H$ reduces to normal form.*

Definition 4.8 *A is a* subformula *of B iff (1) B is A; or (2) B is $B' \to B''$ and A is a subformula of B' or B''; or (3) B is $\Box B'$ and A is a subformula of B'. We say that a derivation $\Gamma, \Delta \vdash x : A$ has the* subformula property *if for all lwffs $y : B$ used in the derivation, B is either a subformula, or the negation of a subformula of some formula in $\{B' \mid z : B' \in \Gamma \cup \{x : A\}\}$. We will sometimes speak loosely of $x : A$ being a subformula of $y : B$, meaning A is a subformula of B.*

Fact 4.9 *If Π is a normal derivation in K or $K + \mathcal{T}_H$, then Π satisfies the subformula property.*

So far, we have considered extensions of K with Horn relational theories. There is, however, no reason why we should not have relational theories that make use of an arbitrary logic. We just have to extend the language and add appropriate rules and axioms. However, *irrespective* of which logic we allow in the labelling algebra, the rules of K dictate that the only way that derivations there can contribute to lwff derivations is via propositions of the form $x\ R\ y$, thus our normalization theorem for K in fact extends to K plus an arbitrary relational theory $\mathcal{T}$.[5] To summarize:

Fact 4.10 *In the logic $K + \mathcal{T}$ the two parts of the proof system are rigorously separated: lwff judgements can depend on rwff judgements, but not vice versa. Thus any normal derivation of an lwff in $K + \mathcal{T}$ is structured as a central derivation in the base logic K 'decorated' with normal subderivations in the relational theory $\mathcal{T}$, which attach onto the central derivation through instances of $\Box E$.*[6]

[5] Normalization and subformula property for derivations of rwffs in a first-order relational theory $\mathcal{T}$ can be easily shown by adapting standard results for first-order logic (cf. [16]).

[6] When $\Diamond$ is added explicitly, then the $\mathcal{T}$-subderivations attach onto the central K-derivation also through instances of $\Diamond I$.

$$
\begin{array}{|c|c|c|}
\hline
\dfrac{\begin{array}{c}[\rho \supset \emptyset] \\ \vdots \\ \emptyset\end{array}}{\rho}\ \emptyset E &
\dfrac{\begin{array}{c}[\rho_1] \\ \vdots \\ \rho_2\end{array}}{\rho_1 \supset \rho_2} \supset I \qquad \dfrac{\rho_1 \supset \rho_2 \quad \rho_1}{\rho_2} \supset E &
\dfrac{\rho}{\forall x(\rho)}\ \forall I \qquad \dfrac{\forall x(\rho)}{\rho[t/x]}\ \forall E \\
\hline
\end{array}
$$

Where, in $\forall I$, x must not occur free in any open assumption on which ρ depends.

Figure 4: The rules of ND_R

This enforced separation between the base logic and the labelling algebra is in the philosophical spirit of LDSs, and it also provides extra structure that is pragmatically useful: since derivations of rwffs use only the resources of the labelling algebra, we may be able to employ theory specific reasoners successfully to automate proof construction. However, in exchange for this extra structure there are limits to the generality of the formulation.

Consider an extension of the labelling algebra to a full first-order theory. To keep distinct the syntax of the base logic from the labelling algebra, we will use connectives from boolean logic — $\emptyset$ (falsum), $\supset$ (implies), $\forall$ — for compound relational formulae in the labelling algebra; as notation, we henceforth assume that the possibly subscripted variable ρ ranges over such rwffs. First-order properties of R are now added as axioms (or rules) directly in their full form, and the *first-order relational theory* $\mathcal{T}_F$ is obtained by extending ND_R (the first-order ND system of R) with a collection C_R of such axioms. For example, for restricted (i, j, m, n) convergency and for irreflexivity we add:

$$\overline{\forall x \forall y \forall z((x\ R^i\ y \cap x\ R^j\ z) \supset \exists u(y\ R^m\ u \cap z\ R^n\ u))}\ rconv \text{ (schematic)}$$

$$\overline{\forall x(\sim (x\ R\ x))}\ irrefl$$

The rules of ND_R are given in Figure 4; rwffs over other connectives (e.g. $\sim$ (negation), $\cap$ (and), $\cup$ (or), $\exists$) and corresponding rules are defined as usual, and we will explicitly use them in the following. We have:

Theorem 4.11 *There are modal logics corresponding to Kripke frames with accessibility relation defined by a collection C_R of first-order axioms which are not correctly represented in $K + \mathcal{T}_F$ with $\mathcal{T}_F = ND_R + C_R$.*

Proof We give an example. According to [21, p.173], the Kripke frame defined by

$$C \equiv \{\ \forall x \forall y \forall z((x\ R\ y \cap x\ R\ z) \supset (y\ R\ z \cup z\ R\ y))\ \}$$

corresponds to the modal logic with axiom schema

$$\neg\Box(\Box A \to B) \to \Box(\Box B \to A)\,.$$

If we assume that A and B are different sentence letters, then a normal proof of this in $K + ND_R + C$ must have the form

$$\dfrac{\dfrac{\dfrac{\begin{array}{c}[x : \neg\Box(\Box A \to B)]^1\ [x\ R\ y]^2\ [y : \Box B]^3\\ \Pi \\ y : A\end{array}}{y : \Box B \to A}\to I^3}{x : \Box(\Box B \to A)}\Box I^2}{x : \neg\Box(\Box A \to B) \to \Box(\Box B \to A)}\to I^1$$

What might Π be? We can use Fact 4.9 to explore all the possibilities. Since A is atomic, Π must end in an application of an elimination rule; by examining the possibilities we see that it must be an application of $\bot E$, since clearly it is not possible to derive $y : A$ directly from the available hypotheses using other elimination rules. Thus the only possible form for Π is

$$\dfrac{\dfrac{[x : \neg\Box(\Box A \to B)]^1 \qquad \dfrac{\dfrac{\dfrac{[y : \Box B]^3 \qquad \begin{array}{c}[x\ R\ y]^2\ [y : \neg A]^4\ [x\ R\ z]^5\ [z : \Box A]^6\\ \Pi_R \\ y\ R\ z\end{array}}{z : B}\Box E}{z : \Box A \to B}\to I^6}{x : \Box(\Box A \to B)}\Box I^5}{x : \bot}\neg E}{y : A}\bot E^4$$

where Π_R is a derivation purely in the relational theory $ND_R + C$. But

$$x\ R\ y, x\ R\ z \nvdash y\ R\ z \text{ in } ND_R + C,$$

so $K + ND_R + C$ cannot prove the characteristic axiom for the frames defined by C, i.e. $K + ND_R + C$ is not complete with respect to the semantics. $\Box$

Clearly, if R were also symmetric, then $x\ R\ y, x\ R\ z \vdash y\ R\ z$. Hence, this particular counter-example to completeness does not hold for extensions of the logic KB, for which, however, other counter-examples can be devised in a similar way. Note also that incompleteness can be shown by means of other modal formulae, but the provability of the corresponding modal axiom is philosophically the first requirement to be fulfilled by the addition of a relational rule. For instance, by similar reasoning, we can show that $x : \Box A \to \Diamond A$ does not follow from $K + ND_R + \{\forall x \exists y (x\ R\ y\}\}$.

4.2 Universal Falsum

The reason for the incompleteness of $K + \mathcal{T}_F$ in the proof of Theorem 4.11 is easy to find; we could imagine replacing Π_R above with

$$\dfrac{\dfrac{\dfrac{[y:\neg A]^4 \qquad \dfrac{[z:\Box A]^6 \qquad \begin{array}{c}[x\ R\ y]^2\ [x\ R\ z]^5\ [y\ R\ z \supset \emptyset]^7\\ \vdots\\ z\ R\ y\end{array}}{y:A}\,\Box E}{y:\bot}\,\neg E}{\emptyset}\,\star}{y\ R\ z}\,\emptyset E^7$$

since we can show that

$$x\ R\ y, x\ R\ z, y\ R\ z \supset \emptyset \vdash z\ R\ y \text{ in } ND_R + C.$$

What we need is some rule $\star$ to allow us to propagate falsum not only between worlds, like gf, but also between the base logic and the relational theory; i.e. collapsing $x:\bot$ and $\emptyset$ together. We can add rules

$$\frac{x:\bot}{\emptyset}\ uf_1 \qquad\qquad \frac{\emptyset}{x:\bot}\ uf_2$$

to K to get the system K^{uf} which has what we call a *universal falsum*. Clearly with universal falsum we loose the separation between the two theories described in Fact 4.10.

Fact 4.12 *In the logic K^{uf} (and, a fortiori, in $K^{uf} + \mathcal{T}_F$) the two parts of the proof system are* not *separated: lwff judgements can depend on rwff judgements, and vice versa.*

In fact, we can show that $K^{uf} + \mathcal{T}_F$, unlike $K + \mathcal{T}$, is essentially equivalent to the usual semantic embedding of modal logics in first-order logic.

Definition 4.13 $(\cdot)^*$ *is a translation of labelled propositional modal logic into first-order logic:*

$$
\begin{array}{rcl@{\qquad}rcl}
(\emptyset)^* & \leadsto & \bot & (x:\bot)^* & \leadsto & \bot \\
(x\ R\ y)^* & \leadsto & R(x,y) & (x:p)^* & \leadsto & P(x) \\
(\rho_1 \supset \rho_2)^* & \leadsto & (\rho_1)^* \rightarrow (\rho_2)^* & (x:A \rightarrow B)^* & \leadsto & (x:A)^* \rightarrow (x:B)^* \\
(\forall x(\rho))^* & \leadsto & \forall x((\rho)^*) & (x:\Box A)^* & \leadsto & \forall y(R(x,y) \rightarrow (y:A)^*) \\
(\Delta)^* & \leadsto & \{(\rho)^* \mid \rho \in \Delta\} & (\Gamma)^* & \leadsto & \{(x:A)^* \mid x:A \in \Gamma\}
\end{array}
$$

Theorem 4.14 *Let C_R be an arbitrary collection of first-order axioms about R, and φ an arbitrary lwff or rwff. We have that $\Gamma, \Delta \vdash \varphi$ in $K^{uf} + ND_R + C_R$ iff $C_R, (\Gamma)^*, (\Delta)^* \vdash (\varphi)^*$ in first-order logic.*

Proof Since reasoning about labels is directly translated, we only treat the case when φ in an lwff. Left to right is simple, since we can find derived rules in first-order logic corresponding to each rule of K^{uf}; e.g.

$$
\dfrac{\begin{array}{c}[x\ R\ y]^1 \\ \vdots \\ y:A\end{array}}{x:\Box A}\ \Box I^1
\qquad \leadsto \qquad
\dfrac{\dfrac{\begin{array}{c}[R(x,y)]^1 \\ \vdots \\ (y:A)^*\end{array}}{R(x,y) \rightarrow (y:A)^*}\ \rightarrow I^1}{\forall y(R(x,y) \rightarrow (y:A)^*)}\ \forall I
\quad [\equiv (x:\Box A)^*]
\tag{5}
$$

The other direction is trickier. However, we know that derivations in first-order logic can be normalized [16, p. 40], thus we can assume Π is a normal derivation of $C_R, (\Gamma)^*, (\Delta)^* \vdash (x:A)^*$, and observe that it is possible to translate this derivation directly into $K^{uf} + ND_R + C_R$; e.g. if we reverse $\leadsto$ in (5), we can see that since a normal derivation of $(x:\Box A)^*$ must have exactly the form (the sequence of introduction rules) given there, and, by induction, the same translation can be performed on the subderivation of $(y:A)^*$ from $[(x\ R\ y)^*]$, it is possible to translate this into a derivation in $K^{uf} + ND_R + C_R$. We can do the same with the elimination rules. All we have to do is, occasionally, insert extra rules translating between falsum for rwffs and falsum for lwffs. $\Box$

Under the assumption (cf., for instance, [12]) that semantic embedding in first-order logic is sound and complete with respect to the appropriate Kripke semantics, we have that:

Corollary 4.15 *$K^{uf} + \mathcal{T}_F$ is sound and complete.*

4.3 Local Falsum

In the rules of K, rwffs interact with lwffs through the $\Box E$ rule and this changes the label of the major premise. But this is not the only rule which changes worlds; $\bot E$, as we have discussed, also has this property. To complete our investigation of alternative formulations, we consider the other end of the spectrum from universal falsum where, by restricting $\bot E$, falsum is local and cannot move arbitrarily between worlds:

$$\cfrac{\begin{array}{c}[x:A \to \bot] \\ \vdots \\ x:\bot\end{array}}{x:A}\ \bot E^{lf}$$

Call K^{lf} the system obtained from K by replacing $\bot E$ with its restricted form $\bot E^{lf}$. Note that in K^{lf} we can propagate $\bot$ forwards indirectly: given $x:\bot$ we have $x:\Box\bot$, and thus $y:\bot$ when $x\ R\ y$; i.e.

$$\cfrac{\cfrac{x:\bot}{x:\Box\bot}\ \bot E^{lf} \qquad x\ R\ y}{y:\bot}\ \Box E$$

But we cannot propagate $\bot$ to an arbitrary world:

Proposition 4.16 *There is no derivation of $y:\bot$ from $x:\bot$ in K^{lf}.*

To show this we prove:

Lemma 4.17 *If there are no applications of $\bot E$ in a derivation in K then normalization of the derivation cannot introduce one.*

Proof By examining the transformations involved in reducing a derivation to normal form. $\Box$

Proof [of Proposition 4.16] Since K^{lf} is a fragment of K, a derivation Π of $y:\bot$ from $x:\bot$ in K^{lf} would have a normal form Π' in K. Since any such derivation needs to make use of $\bot E$, which, by Lemma 4.17, must already be present in the un-normalized form of Π, no such derivation can exist in K^{lf}. $\Box$

In the same way, we can prove that, since gf is not derivable, Proposition 4.2 fails for K^{lf}.

Proposition 4.18 *The connectives $\Box$ and $\Diamond$ are not interdefinable in K^{lf}.*

We need:

Lemma 4.19 *A normal form derived rule in K suitable for the substitution (1) in Proposition 4.2 involves a step application*

$$\dfrac{\begin{array}{c}[x : A \to \bot] \\ \vdots \\ y : \bot\end{array}}{x : A}\ \bot E$$

where we are not able to assume that $y\ R\ x$.

Proof By examination of the possible normal derivations. $\Box$

Proof [of Proposition 4.18] Consider case (1) in the proof of Proposition 4.2. Assume Π is a suitable derivation in K^{lf}, then, since Π is also a derivation in K, it has a normal form Π' in K. However, by Lemmata 4.17 and 4.19 such a derivation in K^{lf} does not exist, since Π', and thus Π, must contain unrestricted applications of $\bot E$. $\Box$

Proposition 4.18 shows that K^{lf} is not in general suitable for formalizing modal logics, since we are not able to propagate falsum to inaccessible worlds. However it is easy to show that in fact we only ever have to deal with worlds accessible in some way from each other. Given, as we have observed, that we can propagate $\bot$ forwards in K^{lf}, if R is symmetrical we also have a backwards propagation:

$$\dfrac{\dfrac{x : \bot}{x : \Box\bot}\ \bot E^{lf} \quad \dfrac{y\ R\ x}{x\ R\ y}\ R\_symm}{y : \bot}\ \Box E$$

Thus K^{lf} can be used to formalize certain logics after a fashion (if the relational theory $\mathcal{T}_F$ is inconsistent or if R is *universal*, so that $x\ R\ y$ for all x, y, then we get this much more simply).[7] However the resulting formalization is fundamentally unsatisfactory, since it lacks important metatheoretic properties that we get in K; namely, we have:

[7] Given that $S5$ is correct with respect to the class of universal frames [4, p.178], it is possible to prove that $\Gamma, \Delta \vdash x : A$ in $KT5$ iff $\Gamma, \Delta \vdash x : A$ in $K^{lf}T5$, since, when R is universal, $\Box$ and $\Diamond$ are interdefinable, and $\bot E$ and $\bot E^{lf}$ are interderivable (but the derivations are not normal).

Theorem 4.20 *Derivations in K^{lf} do not have normal forms satisfying the subformula property.*

Proof As we observed above, there is a derivation of $y : \perp$ from $x\ R\ y$ and $x : \perp$ in K^{lf}. However, there cannot be a normal one satisfying the subformula property. $\square$

5 Implementation and its Correctness

5.1 Implementation

We have used Paulson's Isabelle system [15] to implement and interactively construct derivations with the modal logics we presented. The logical basis of Isabelle is a natural deduction presentation of minimal implicational predicate logic with universal quantification over all higher-types [14].[8] We call this metalogic $\mathcal{M}$; to prevent object/meta confusion we use $\bigwedge$ to represent Isabelle's universal quantifier and $\Rightarrow$ for implication.

An object logic is encoded in Isabelle by declaring a theory, which consists of a signature and axioms, which are formulae in the language of $\mathcal{M}$. The axioms are used to establish the validity of judgements, which are assertions about syntactic objects declared in the signature [10]. Derivations are constructed by deduction in the metalogic.

In our work, we declare a theory $\mathcal{M}_K$, which encodes K. The signature of $\mathcal{M}_K$ declares two types *label* and o, which denote labels and unlabelled modal formulae, respectively. Connectives and modal operators are declared as typed constants over this signature, i.e. *box* of type $o \Rightarrow o$. There are two judgements, which correspond to predicate symbols in the metalogic: $\mathcal{L}$ and $\mathcal{A}$, which stand for 'Labelled Formula' and 'Accessibility'. $\mathcal{L}(x : A)$ and $\mathcal{A}(x\ R\ y)$ respectively express the judgements that $x : A$ is a provable lwff and that $x\ R\ y$ is a provable rwff. The axioms for $\mathcal{L}$ are a direct axiomatization of the rules in Figure 1.

Figure 5 contains our entire Isabelle declaration for the theory $\mathcal{M}_K$. Some brief explanations are in order (further details on Isabelle syntax and

[8] Isabelle's logic also contains equality (that of the λ-calculus under α, β, and η-conversion), but we do not need to consider this, since, in the analysis of derivations in the metalogic, we shall identify terms with their $\beta\eta$ normal forms. This is possible as terms in our metatheories are terms in the simply-typed λ-calculus (with additional function constants) and every term can be reduced to a normal form that is unique up to α-conversion.

```
K = Pure +
types (* Definition of type constructors *)
  label,o 0
arities (* Addition of the arity 'logic' to the existing types *)
  label, o :: logic
consts  (* Logical Connectives and Judgements L and A) *)
  False        :: "o"
  -->          :: "[o, o] => o"               (infixr 25)
  box          :: "o => o"                    ("[]_" [50] 50)
  dia          :: "o => o"                    ("<>_" [50] 50)
  L            :: "[label, o] => prop"        ("(_ : _)" [0,0] 100)
  A            :: "[label, label] => prop"    ("(_ R _)" [0,0] 100)
rules (* Axioms representing the object-level rules *)
  FalseE       "(x:A --> False ==> y: False) ==> x:A"
  impI         "(x:A ==> x:B) ==> x:A --> B"
  impE         "x:A ==> x:A --> B ==> x:B"
  boxI         "(!!y. (x R y ==> y:A)) ==> x:[]A"
  boxE         "x:[]A ==> x R y ==> y:A"
  diaI         "y:A ==> x R y ==> x:<>A"
  diaE         "x:<>A ==> (!!y. y:A ==> x R y ==> z:B) ==> z:B"
end
```

Figure 5: Isabelle Encoding of K

theory declarations can be found in [15]). First, we shall use `typewriter font` for displaying concrete Isabelle syntax which has come from actual Isabelle sessions. `Pure` encodes Isabelle's metalogic $\mathcal{M}$. The operators `!!` and `==>` are concrete syntax in Isabelle for universal quantification ($\bigwedge$) and implication ($\Rightarrow$) in $\mathcal{M}$ respectively. The use of mixfix operators, declared with information for Isabelle's parser, allows us to abbreviate `box` with `[]`, `dia` with `<>`, `L(x:A)` with `x:A`, and `A(xRy)` with `xRy`. Note that in axioms, free variables are implicitly outermost universally quantified. Finally, comments are added between '`(*`' and '`*)`'.

Logics $L = K+\mathcal{T}$ are formed by extending $\mathcal{M}_K$ with appropriate theories $\mathcal{M}_\mathcal{T}$, which encode $\mathcal{T}$. The axioms for $\mathcal{A}$ are given by directly translating Horn relational rules to axioms in $\mathcal{M}$: each rule corresponds to an iterated (Curried) implication where the assumptions of the rule together imply the conclusion.

Theories in Isabelle correspond to instances of an abstract datatype in the ML programming language and Isabelle provides means for creating elements of these types, extending them, and combining them. We use these facilities to combine and extend our modal theories. This is best illustrated by an example. `KT` is obtained by extending `K` with the axiom `R_refl`; this is specified as follows.

```
KT = K +
rules
  R_refl        "x R x"
end
```

Again, recall that outermost quantifiers are left implicit, so the above is shorthand for adding `!! x. x R x` as an axiom to `K`. Similarly, `K4` is formed by extending `K` with `R_trans`.

```
K4 = K +
rules
  R_trans       "x R y ==> y R z ==> x R z"
end
```

We may now obtain `KT4`, i.e. `S4`, by similarly extending `KT` (or `K4` or `K`). Alternatively, we may apply an ML-function `merge_theories` to `KT` and `K4`. As remarked above, `KT4` inherits theorems and derived rules from its ancestor logics. As an example, consider the `KT4`-theorem `x:[]A <-> [][]A`. `x:[]A --> [][]A` and `x:[][]A --> []A` are theorems of `K4` and `KT`, respectively:

$$\dfrac{\dfrac{\dfrac{\dfrac{[x\!:\!\Box A]^3 \qquad \dfrac{[x\ R\ y]^2 \quad [y\ R\ z]^1}{x\ R\ z}\,R\_trans}{z\!:\!A}\,\Box E}{y\!:\!\Box A}\,\Box I^1}{x\!:\!\Box\Box A}\,\Box I^2}{x\!:\!\Box A \to \Box\Box A}\,\to I^3 \qquad\qquad \dfrac{\dfrac{[x\!:\!\Box\Box A]^1 \quad \dfrac{}{x\ R\ x}\,R\_refl}{x\!:\!\Box A}\,\Box E}{x\!:\!\Box\Box A \to \Box A}\,\to I^1 \tag{6}$$

In Appendix A, we show how these theorems are interactively proved in Isabelle in their corresponding theories and then applied to show that the following equivalence is a theorem of KT4.

$$\dfrac{x\!:\!\Box A \to \Box\Box A \quad x\!:\!\Box\Box A \to \Box A}{x\!:\!\Box A \leftrightarrow \Box\Box A}\,\leftrightarrow I$$

(Note that this requires adding a definition of $\leftrightarrow$ to our theory, which can be done in the standard way.)

As a further example of theory definition, `K2` is obtained by extending `K` with the constant function symbol `g` and with the axioms `R_conv1` and `R_conv2`:

```
K2 = K +
consts
  g           :: "[label,label,label] => label"
rules
  R_conv1     "x R y ==> x R z ==> y R g(x,y,z)"
  R_conv2     "x R y ==> x R z ==> z R g(x,y,z)"
end
```

In the appendix we use this theory to prove $x\!:\!\Diamond\Box A \to \Box\Diamond A$, (see the proof in Section 2.4), which is `K2`'s characteristic axiom. The examples we work through in Isabelle should help convince the reader that the approach we have taken to interactive theorem proving for modal logics is both simple and flexible. In particular, it supports the hierarchical structuring of theories and inheritance of theorems between them.

5.2 Correctness

When one logic encodes another, correctness of the encoding must be shown. A technique established with the Edinburgh LF [10] is to demonstrate a correspondence between derivations in the object-logic and derivations in the metalogic by considering certain normal forms for derivations in the metalogic. In what follows, we write $\mathcal{L}(\Gamma)$ and $\mathcal{A}(\Delta)$ for the sets $\{\mathcal{L}(x_1 : A_1), \ldots, \mathcal{L}(x_n : A_n)\}$ and $\{\mathcal{A}(x_1\ R\ y_1), \ldots, \mathcal{A}(x_m\ R\ y_m)\}$.

Definition 5.1 *$\mathcal{M}_L$ is* faithful *(with respect to L) iff (1) $\mathcal{L}(\Gamma), \mathcal{A}(\Delta) \vdash_{\mathcal{M}_L} \mathcal{L}(x : A)$ implies $\Gamma, \Delta \vdash_L x : A$, and (2) $\mathcal{L}(\Gamma), \mathcal{A}(\Delta) \vdash_{\mathcal{M}_L} \mathcal{A}(x\ R\ y)$ implies $\Gamma, \Delta \vdash_L x\ R\ y$. $\mathcal{M}_L$ is* adequate *(with respect to L) iff the converses of (1) and (2) hold.*

Lemma 5.2 *$\mathcal{M}_L$ is faithful.*

Proof Following Prawitz, call a *thread* a sequence of formulae in a derivation tree leading from some assumption to the root. A *branch* in a derivation is the initial segment of a thread ending at either the first minor premise of a $\rightarrow E$ rule encountered, or the conclusion of the derivation if no such minor premise occurs. We use the fact [16] that derivations in $\mathcal{M}_L$ have an *expanded normal form* in which there are no maximal formulas and each branch leads to a *minimum formula* of the form $\mathcal{L}(x : A)$ or $\mathcal{A}(x\ R\ y)$.

The proof proceeds by induction on the size of the expanded normal form of $\mathcal{M}_L$-derivations of $\mathcal{L}(x : A)$ and of $\mathcal{A}(x\ R\ y)$ from $\mathcal{L}(\Gamma)$ and $\mathcal{A}(\Delta)$. In the base case, if $\mathcal{L}(x : A)$ follows from an assumption in $\mathcal{L}(\Gamma)$, then $x : A$ is an assumption in Γ, so trivially $\Gamma, \Delta \vdash_L x : A$. The situation is similar for a proof of $\mathcal{A}(x\ R\ y)$ from an assumption in $\mathcal{A}(\Delta)$.

In the step case, a branch begins with an axiom followed by a sequence of elimination rules. We proceed by showing that the application of each axiom in $\mathcal{M}_L$ corresponds to an object level inference in L. All of the cases are simple and we give two representative cases below: the axiom `boxI` from $\mathcal{M}_K$ and a Horn axiom from $\mathcal{M}_{\mathcal{T}}$.

In the case of `boxI`, let $x : A$ be $z : \Box B$ for some z and B. The $\mathcal{M}_L$-derivation must have the structure shown at the top of Figure 6, where $\bigwedge E^*$ stands for two consecutive applications of $\bigwedge E$. It contains an $\mathcal{M}_L$-derivation of $\bigwedge y(\mathcal{A}(z\ R\ y) \Rightarrow \mathcal{L}(y : B))$ from $\mathcal{L}(\Gamma)$ and $\mathcal{A}(\Delta)$, which, by expanded normal form, consists of an $\mathcal{M}_L$-derivation of $\mathcal{L}(y : B)$ from $\mathcal{L}(\Gamma)$ and $\mathcal{A}(\Delta \cup \{z\ R\ y\})$, where y is not free in the assumptions, followed first by a $\Rightarrow I$, discharging the assumption $\mathcal{A}(x\ R\ y)$, and then by a $\bigwedge I$. An L-derivation of $y : B$ from Γ and $\Delta \cup \{z\ R\ y\}$, where y is not free in the assumptions, is given by the induction hypothesis. Applying $\Box I$ gives an L-derivation of $z : \Box B$ from Γ and Δ.

Alternatively, consider a Horn axiom which is part of the relational theory corresponding to $\mathcal{M}_{\mathcal{T}}$. The $\mathcal{M}_L$-derivation must comprise a sequence

$$
\dfrac{
\dfrac{\begin{array}{c}\bigwedge x \bigwedge A((\bigwedge y(\mathcal{A}(x\ R\ y) \Rightarrow \\ \mathcal{L}(y\!:\!A))) \Rightarrow \mathcal{L}(x\!:\!\Box A))\end{array}}{\bigwedge y(\mathcal{A}(z\ R\ y) \Rightarrow \mathcal{L}(y\!:\!B)) \Rightarrow \mathcal{L}(z\!:\!\Box B)}\ \bigwedge E^{*}
\qquad
\dfrac{\dfrac{\begin{array}{c}[\mathcal{A}(z\ R\ y)]^{1} \\ \vdots \\ \mathcal{L}(y\!:\!B)\end{array}}{\mathcal{A}(z\ R\ y) \Rightarrow \mathcal{L}(y\!:\!B)}\ \Rightarrow I^{1}}{\bigwedge y(\mathcal{A}(z\ R\ y) \Rightarrow \mathcal{L}(y\!:\!B))}\ \bigwedge I
}{\mathcal{L}(z\!:\!\Box B)}\ \Rightarrow E
$$

$$
\dfrac{
\dfrac{
\dfrac{\begin{array}{c}\bigwedge x \bigwedge y \bigwedge z(\mathcal{A}(x\ R\ y) \Rightarrow \\ (\mathcal{A}(x\ R\ z) \Rightarrow \mathcal{A}(y\ R\ g(x,y,z))))\end{array}}{\mathcal{A}(u\ R\ v) \Rightarrow (\mathcal{A}(u\ R\ w) \Rightarrow \mathcal{A}(v\ R\ g(u,v,w)))}\ \bigwedge E^{*}
\quad
\begin{array}{c}\vdots \\ \mathcal{A}(u\ R\ v)\end{array}
}{\mathcal{A}(u\ R\ w) \Rightarrow \mathcal{A}(v\ R\ g(u,v,w))}\ \Rightarrow E
\quad
\begin{array}{c}\vdots \\ \mathcal{A}(u\ R\ w)\end{array}
}{\mathcal{A}(v\ R\ g(u,v,w))}\ \Rightarrow E
$$

Figure 6: The metalevel derivations formalizing $\Box I$ and R_conv1

of $\bigwedge E$ steps, one for each quantifier, followed by a sequence of $\Rightarrow E$ steps, one for each premise. For concreteness, consider the axiom R_conv1, where $x\ R\ y$ is $v\ R\ g(u, v, w)$ for some u, v, w. The $\mathcal{M}_L$-derivation must have the structure shown at the bottom of Figure 6, where $\bigwedge E^*$ stands for three consecutive applications of $\bigwedge E$. L-derivations of $u\ R\ v$ and $u\ R\ w$ from Γ and Δ are given by induction hypotheses. Applying R_conv1 gives an L-derivation of $v\ R\ g(u, v, w)$ from Γ and Δ. $\Box$

Lemma 5.3 *$\mathcal{M}_L$ is adequate.*

Proof By induction on the structure of the L-derivations of $x : A$ and of $x\ R\ y$ from Γ and Δ. The base cases are trivial, and we treat only the step cases.

First, we consider the propositional and the modal rules (i.e. the rules of K) individually. For example, for $\Box I$, let $x : A$ be $z : \Box B$, and $\Box I$ is applied to an L-derivation of $y : B$ from Γ and $\Delta \cup \{z\ R\ y\}$, where y is not free in the assumptions. An $\mathcal{M}_L$-derivation of $\mathcal{L}(y : B)$ from $\mathcal{L}(\Gamma)$ and $\mathcal{A}(\Delta \cup \{z\ R\ y\})$, where y is not free in the assumptions, i.e. an $\mathcal{M}_L$-derivation of $\bigwedge y(\mathcal{A}(z\ R\ y) \Rightarrow \mathcal{L}(y : B))$ from $\mathcal{L}(\Gamma)$ and $\mathcal{A}(\Delta)$, is given by induction hypothesis. Conclude by building an $\mathcal{M}_L$-derivation like that at the top of Figure 6.

In second case, a relational rule has been applied. Consider the case of R_conv1. $x\ R\ y$ is $v\ R\ g(u, v, w)$, and R_conv1 is applied to L-derivations of $u\ R\ v$ and $u\ R\ w$ from Γ and Δ. $\mathcal{M}_L$-derivations of $\mathcal{A}(u\ R\ v)$ and $\mathcal{A}(u\ R\ w)$ from $\mathcal{L}(\Gamma)$ and $\mathcal{A}(\Delta)$ are given by induction hypotheses. Conclude by building an $\mathcal{M}_L$-derivation like that at the bottom of Figure 6. $\Box$

By Lemma 5.2 and Lemma 5.3 we have that:

Theorem 5.4 *$\mathcal{M}_L$ is faithful and adequate.*

6 Related Work

Our work combines an LDS presentation of modal logics with a logical framework to provide a natural deduction presentation of modal logics in a uniform way based on their semantics. Here we compare this with related work in natural deduction, Labelled Deductive Systems, and semantic embedding.

6.1 Natural Deduction

Prawitz [16] discusses a rule for necessitation ($\Box$) introduction in $S4$ and $S5$ with the 'non-local' side condition that all the supporting assumptions are modal (i.e. the main connective is $\Box$), in the case of $S4$, or modal formulae and their negation, in the case of $S5$. However, such a rule cannot be formalized by a pure proof rule, e.g. one that may be applied in any context of assumptions; hence it cannot be directly encoded within a logical framework. A solution to this problem is given, as mentioned earlier, in [2, §4.4], where the proof system is factored into two ordinary pure single-conclusioned consequence relations. Unfortunately, the result is far removed from the standard presentations based on accessibility relations or characteristic axioms. Also there is no attempt to modularize structure or correctness: only a particular modal logic is analyzed and it is not apparent how to generalize the results in a uniform way.

Another approach to the formalization of 'non-local' conditions in a logical framework is to manage assumptions explicitly with sequents, e.g. [7, 24]. The Isabelle system distribution contains such an encoding due to Martin Coen which uses several auxiliary judgements to give complex encodings of T, $S4$, and $S4.3$. Similar problems would result from trying to formalize directly the kind of prefixed tableaux systems suggested, for example, by Fitting [7].

6.2 Labelled Deductive Systems

Our work is inspired by the LDS approach proposed by Gabbay, and further developed for modal logics, in parallel with our work, by Russo [18]. Gabbay introduces LDSs as a general and unifying methodology for presenting almost any logic [8]. To support this generality his LDS metatheory and presentations are based on a notion of diagrams and logic data-bases, which are manipulated by rules with multiple premises and conclusions. For example [8, p.57] presents the rule for $\Diamond E$ as

$$\frac{s\!:\!\Diamond B}{\textit{create } r,\ s < r \textit{ and } r\!:\!B}$$

the application of which updates a modal data-base with the two new conclusions (a rule to the same effect is given in [18]). The formal details are quite different from our proposal, where the rule for $\Diamond E$ given in Figure 1 is represented in the metalevel of Isabelle by the following axiom, which directly formalizes a natural deduction rule:

$$\bigwedge x \bigwedge z \bigwedge A \bigwedge B(\mathcal{L}(x\!:\!\Diamond A) \Rightarrow (\bigwedge y(\mathcal{L}(y\!:\!A) \Rightarrow \mathcal{A}(x\ R\ y)) \Rightarrow \mathcal{L}(z\!:\!B)) \Rightarrow \mathcal{L}(z\!:\!B))\,.$$

There is another difference between our work and theirs that is worth emphasizing. In our work, we have identified an important property of the structured presentation of logics, their combination, and extension. Namely, there is tension between modularity and extensibility: a narrow interface between the base logic and labelling algebra provides a better (more modular) metatheory, but can limit the degree to which we can make use of extensions to the labelling algebra. In our approach, the use of a metalogic with different judgements serves to separate the base logic and the labelling algebra. This separation is critical: it is only when we attempt to modularize and separate these two theories formally and define a precise interface between them that we see that only limited modularity (i.e. there are limits to the relational theories) is actually possible.

Of course, in implementing particular LDSs Gabbay and Russo can similarly separate theories. The precise nature of this would be reflected in the rules they choose for propagating results between data-bases. It should be the case that if their rules enforce a similar separation, then they will encounter similar limitations to those reported here. That is, the problems we identify have some generality and should appear in other frameworks where theories are separated and results are communicated in a limited way between them.

The kind of labelled natural deduction encoding we employ is closest to the work of Simpson [20]. However his focus, proof techniques, and applications are based on using LDSs to investigate intuitionistic versions of modal logics, and his correctness considerations are quite different. Moreover, his relations have no independent theory with which one can work.

Note that the universal falsum approach is adopted explicitly in [18]. Simpson's approach is different, and difficult to compare: he treats rwffs only as assumptions in inferences of lwffs via his 'geometric' rules, which are derivable in our systems. An example of an approach in which, like with local falsum, local inconsistency does not imply global inconsistency, is the work of Giunchiglia and Serafini [9], who show that particular 'multicontext systems', where (indexed) formulae are translated between contexts using 'bridge rules', define the same classes of provable formulae as certain standard modal logics. However their approach is, in general, radically different from ours, and not comparable.

6.3 Translation and Semantic Embedding

We conclude by mentioning work on translating modal logics into first-order logics, e.g. [12, 13]. As sketched in Definition 4.13, these approaches typically label all subformulae with worlds and combine the modal and relational theory in a theory suitable for standard first-order provers. The emphasis is on automatic, but not necessarily 'natural', theorem proving. Moreover, by design, there is no separation between the relational theory, any kind of base modal theory, and first-order logic itself; i.e. there is precisely one falsum from which one can conclude arbitrary relational or labelled formulae.

7 Conclusions

We have given a modular presentation and correctness proofs for implementing a large and well-known class of propositional modal logics in the Isabelle logical framework. Our approach is based on relational theories comprised of (Horn clause) axioms formalizing the accessibility of worlds in Kripke frames, and it demonstrates, we think, that they fit particularly well into the logical framework setting, capture a large class of standardly considered propositional modal logics, and have pleasant metatheoretic properties (e.g. one can use induction on their structure to show faithfulness and adequacy across an infinite set of extensions). We may use similar techniques to present quantified modal logics, which will be dealt with in a forthcoming companion paper [3].

Our work has also identified an important property of the structured presentation of logics, their combination, and extension. Namely, there is tension between modularity and extensibility: a narrow interface between the base logic and labelling algebra can limit the degree to which we can make use of extensions to the labelling algebra. As a consequence, there are important design decisions in implementing LDSs whose resolution requires predicting the range of possible applications.

References

[1] A. Avron. Simple Consequence Relations. *Information and Computation*, 92:105–139, 1991.

[2] A. Avron, F. Honsell, I. Mason, and R. Pollack. Using Typed Lambda Calculus to Implement Formal Systems on a Machine. *Journal of Automated Reasoning*, 9:309–352, 1992.

[3] D. Basin, S. Matthews, and L. Viganò. Labelled Quantified Modal Logics. Submitted.

[4] B. Chellas. *Modal Logic*. Cambridge University Press, New York, 1980.

[5] M. D'Agostino and D. Gabbay. A generalization of analytic deduction via labelled deductive systems. part I : Basic substructural logics. *Journal of Automated Reasoning*, 13:243–281, 1994.

[6] A. Felty and D. A. Miller. Specifying Theorem Provers in a Higher-Order Logic Programming Language. In E. Lusk and R. Overbeek, editors, *CADE 9*, Berlin, 1988. Springer LNCS 310.

[7] M. Fitting. *Proof Methods for Modal and Intuitionistic Logics*. Kluwer, Dordrecht, 1983.

[8] D. Gabbay. LDS - Labelled Deductive Systems, Volume 1 - Foundations. Technical report, MPI für Informatik, Saarbrücken, 1994.

[9] F. Giunchiglia and L. Serafini. Multilanguage Hierarchical Logics (or: how we can do without modal logics). *Artificial Intelligence*, 65:29–70, 1994.

[10] R. Harper, F. Honsell, and G. Plotkin. A framework for defining logics. *Journal of the ACM*, 40(1):143–184, 1993.

[11] R. P. Nederpelt, J. H. Geuvers, and R. C. de Vrijer, editors. *Selected Papers on Automath.* Elsevier, Amsterdam, 1994.

[12] H. J. Ohlbach. *A Resolution Calculus for Modal Logics.* PhD thesis, Universität Kaiserslautern, Kaiserslautern, Germany, 1988.

[13] H.-J. Ohlbach. Translation methods for non-classical logics: an overview. In *Bulletin of the IGPL*, volume 1, Saarbrücken, 1993.

[14] L. Paulson. The Foundation of a Generic Theorem Prover. *Journal of Automated Reasoning*, 5:363–397, 1989.

[15] L. Paulson. *Isabelle: A Generic Theorem Prover.* LNCS-828. Springer, Berlin, 1994.

[16] D. Prawitz. *Natural Deduction, a Proof-Theoretical Study.* Almqvist and Wiksell, Stockholm, 1965.

[17] D. Prawitz. Ideas and Results in Proof Theory. In J. E. Fensted, editor, *Proc. 2nd Scandinavian Logic Symp.*, Amsterdam, 1971. North-Holland.

[18] A. Russo. Modal Labelled Deductive Systems. Technical Report 95/7, Department of Computing, Imperial College, London, UK, 1995.

[19] J. R. Shoenfield. *Mathematical Logic.* AddisonWesley, 1967.

[20] A. Simpson. *The Proof Theory and Semantics of Intuitionistic Modal Logic.* PhD thesis, University of Edinburgh, Edinburgh, 1993.

[21] J. van Benthem. Correspondence Theory. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic II.* D. Reidel Publishing Company, 1984.

[22] J. van Benthem. *Modal Logic and Classical Logic.* Bibliopolis, Napoli, 1985.

[23] D. van Dalen. *Logic and Structure.* Springer, Berlin, 1994.

[24] L. Wallen. *Automated Deduction in Non-Classical Logics.* MIT Press, Cambridge, Mass., 1990.

A Isabelle Proof Session

In this appendix we illustrate Isabelle proofs for the examples sketched in Section 5.1. Some brief background is required; see [15] for a full account.

Background

Isabelle manipulates *rules*. A rule is a formula

```
!! v1 ... vm. A1 ==>  ... ==> (An ==> A)
```

which is also displayed as follows:

```
!! v1 ... vm. [| A1;  ...; An|] ==> A
```

Rules represent proof states where `A` is the goal to be established and the `Ai` are the subgoals to be proved. Under this view, an initial proof state has the form `A ==> A`, i.e. it has one subgoal, namely `A`. The final proof state *is itself* the desired theorem. Isabelle supports proof construction through higher-order resolution which is roughly analogous to resolution in Prolog. That is, given a proof state with subgoal `B` and a rule as above, then (treating the `vi` as variables for unification) we higher-order unify `A` with `B`. If this succeeds, then the unification yields a substitution `s` and the proof state is updated replacing `B` with the subgoals `s(A1),...,s(An)`. This resolution step can be justified by a sequence of proof steps in the metalogic. Although rules are formalized in a natural deduction style, they may be read as intuitionistic sequents where the `Ai` are the hypotheses. Isabelle has procedures which apply rules in a way that maintains this 'illusion' of working with sequents.

Derivations

We now work through the examples given in Section 5.1. To prove the equivalence of $\Box A$ and $\Box\Box A$ in $S4$ we begin by proving the left-to-right direction in the subtheory $K4$. Our proof corresponds to the first proof-tree given in (6), read bottom up; the following proof is taken verbatim from an Isabelle session with the exception of minor pretty-printing and omission of diagnostic output. We begin with the desired goal.

```
> goal K4.thy "x:[]A --> [][]A";
x : []A --> [][]A
 1. x : []A --> [][]A
```

On the first line state the theory we are using and the theorem to be proved. Isabelle responds with the next 2 lines which give the goal to be proved, and what subgoals must be established to prove it. We proceed by applying our rule for implication introduction `impI`, which was declared in Figure 5. The command `br` directs Isabelle to apply this using resolution to the first subgoal. Isabelle responds with the new subgoal.

```
> br impI 1;
x : []A --> [][]A
 1. x : []A ==> x : [][]A
```

If we read the proof state as a sequent, we must now show `x : [][]A` under the assumption `x : []A`. We proceed with two applications of `boxI`, each of which gives us new relational assumptions, followed by `boxE`:

```
> br boxI 1;
x : []A --> [][]A
 1. !!y. [| x : []A; x R y |] ==> y : []A

> br boxI 1;
x : []A --> [][]A
 1. !!y ya. [| x : []A; x R y; y R ya |] ==> ya : A

> be boxE 1;
x : []A imp [][]A
 1. !!y ya. [| x R y; y R ya |] ==> x R ya
```

The theory $K4$ extends K with the transitivity of R. We apply transitivity using the command `be` to unify one of its assumptions against an assumption in our subgoal.

```
> be R_trans 1;
x : []A --> [][]A
 1. !!y ya. y R ya ==> y R ya
```

This leaves only one remaining goal, which is proved by assumption (`ba`).

```
> ba 1;
x : []A --> [][]A
No subgoals!
```

We can now name this theorem (`LeftToRight`) and use it in subsequent proofs (Isabelle provides *unknowns*, written with a `?` prefix, that may be instantiated later during unification).

```
> val LeftToRight = result();
val LeftToRight = "?x : []?A --> [][]?A"
```

The proof of the converse direction in the theory KT directly mirrors the second proof-tree in (6); we give it here without further comment.

```
> goal KT.thy "x:[][]A --> []A";
x : [][]A --> []A
 1. x : [][]A --> []A

> br impI 1;
Level 1
x : [][]A --> []A
 1. x : [][]A ==> x : []A

> be boxE 1;
x : [][]A --> []A
 1. x R  x

> br R_refl 1;
x : [][]A --> []A
No subgoals!

> val RightToLeft = result();
val RightToLeft = "?x : []?A --> ?A"
```

Having proved both directions, we may now combine them to prove the equivalence in $KT4$.

```
> goal KT4.thy "x:[]A <-> [][]A";
x : []A <-> [][]A
 1. x : []A <-> [][]A

> br iffI 1;
x : []A <-> [][]A
 1. x : []A --> [][]A
 2. x : [][]A --> []A
```

```
> br LeftToRight 1;
x : []A <-> [][]A
 1. x : [][]A --> []A

> br RightToLeft 1;
x : []A <-> [][]A
No subgoals!
```

A final example is the derivation of the characteristic axiom for $K2$ based on the extension of K given in Section 5.1. The proof directly follows that given in Section 2.4.

```
> goal K2.thy "x: <>[]A --> []<>A";
x : <>[]A --> []<>A
 1. x : <>[]A --> []<>A

> br impI 1;
x : <>[]A --> []<>A
 1. x : <>[]A ==> x : []<>A

> br boxI 1;
x : <>[]A --> []<>A
 1. !!y. [| x : <>[]A; x R y |] ==> y : <>A

> be diaE 1;
x : <>[]A --> []<>A
 1. !!y ya. [| x R y; ya : []A; x R ya |] ==> y : <>A

> br diaI 1;
x : <>[]A --> []<>A
 1. !!y ya. [| x R y; ya : []A; x R ya |] ==> ?y3(y, ya) : A
 2. !!y ya. [| x R y; ya : []A; x R ya |] ==> y R ?y3(y, ya)

> be boxE 1;
x : <>[]A --> []<>A
 1. !!y ya. [| x R y; x R ya |] ==> ya R ?y3(y, ya)
 2. !!y ya. [| x R y; ya : []A; x R ya |] ==> y R ?y3(y, ya)

> be R_conv2 1;
```

```
x : <>[]A --> []<>A
 1. !!y ya. x R ya ==> x R ya
 2. !!y ya. [| x R y; ya : []A; x R ya |] ==> y R g(x, y, ya)

> ba 1;
x : <>[]A --> []<>A
 1. !!y ya. [| x R y; ya : []A; x R ya |] ==> y R g(x, y, ya)

> be R_conv1 1;
x : <>[]A --> []<>A
 1. !!y ya. [| ya : []A; x R ya |] ==> x R ya

> ba 1;
x : <>[]A --> []<>A
No subgoals!
```

I N F O R M A T I K

Below you find a list of the most recent technical reports of the research group *Logic of Programming* at the Max-Planck-Institut für Informatik. They are available by anonymous ftp from our ftp server `ftp.mpi-sb.mpg.de` under the directory `pub/papers/reports`. Most of the reports are also accessible via WWW using the URL `http://www.mpi-sb.mpg.de`. If you have any questions concerning ftp or WWW access, please contact `reports@mpi-sb.mpg.de`. Paper copies (which are not necessarily free of charge) can be ordered either by regular mail or by e-mail at the address below.

Max-Planck-Institut für Informatik
Library
attn. Regina Kraemer
Im Stadtwald
D-66123 Saarbrücken
GERMANY
e-mail: `kraemer@mpi-sb.mpg.de`

MPI-I-96-2-003	H. Baumeister	Using Algebraic Specification Languages for Model-Oriented Specifications
MPI-I-96-2-002	D. Basin, S. Matthews, L. Vigano	Labelled Propositional Modal Logics: Theory and Practice
MPI-I-96-2-001	H. Ganzinger, U. Waldmann	Theorem Proving in Cancellative Abelian Monoids
MPI-I-95-011	P. Mutzel	Automatisiertes Zeichnen von Diagrammen
MPI-I-95-010	C. Rüb	On the Average Running Time of Odd-Even Merge Sort
MPI-I-95-009	J. Könemann, C. Schmitz, C. Schwarz	Radix heaps, an efficient implementation for priority queues
MPI-I-95-008	H. J. Ohlbach, R. A.Schmidt, U. Hustadt	Translating Graded Modalities into Predicate Logic
MPI-I-95-007	J. Radhakrishnan, S. Saluja	Lecture Notes Interactive Proof Systems
MPI-I-95-006	95, P. G. Bradford, R. Fleischer, M. Smid	95 A Polylog-Time and $O(n\sqrt{\lg n})$-Work Parallel Algorithm for finding the Row Minima in Totally Monotone Matrices
MPI-I-95-005	J.-H. Hoepmann, M. Papatriantafilou, P. Tsigas	Towards Self-Stabilizing Wait-Free Shared Memory Objects
MPI-I-95-003	P. G. Bradford, R. Fleischer	Matching nuts and bolts faster
MPI-I-95-002	S. Näher, C. Uhrig	LEDA user manual (Version R 3.2)
MPI-I-95-001	M. Smid, C. Thiel, F. Follert, E. Schömer, J. Sellen	Computing a largest empty anchored cylinder, and related problems
MPI-I-94-261	P. Barth, A. Bockmayr	Finite Domain and Cutting Plane Techniques in CLP($\mathcal{PB}$)
MPI-I-94-257	S. Vorobyov	Structural Decidable Extensions of Bounded Quantification
MPI-I-94-254	P. Madden	Report and abstract not published
MPI-I-94-252	P. Madden	A Survey of Program Transformation With Special Reference to *Unfold/Fold* Style Program Development
MPI-I-94-251	P. Graf	Substitution Tree Indexing