Implementing Connection Calculi for First-order Modal Logics

Jens Otten

Institut für Informatik, University of Potsdam
August-Bebel-Str. 89, 14482 Potsdam-Babelsberg, Germany
jeotten@cs.uni-potsdam.de

Abstract

This paper presents an implementation of an automated theorem prover for first-order modal logic that works for the constant, cumulative, and varying domain of the modal logics D, T, S4, and S5. It is based on the connection calculus for classical logic and uses prefixes representing world paths and a prefix unification algorithm to capture the restrictions given by the Kripke semantics of the standard modal logics. This permits a modular and elegant treatment of the considered modal logics and yields an efficient implementation. Details of the calculus, the implementation and performance results on the QMLTP problem library are presented.

1 Introduction

Modal logics add the unary modal operators □ and ◊ to classical logic. They are used to represent the modalities "it is necessarily true that" and "it is possibly true that", respectively. The Kripke semantics of the standard modal logics is defined by a set of worlds and a binary accessibility relation between these worlds. In each single world the classical semantics applies to the standard connectives ¬, ∧, ∨, ⇒, and the modal operators are interpreted with respect to accessible worlds. The properties of the accessibility relation specify the particular modal logic. In first-order modal logics domains of objects are associated with each world, and the standard universal and existential quantifiers are added. First-order modal logics have many applications, e.g. in planning, natural language processing, and program verification. Many of these application require tools for automated reasoning in order to automate the proof search. Whereas there exist a few automated theorem proving systems for some propositional modal logics, e.g. modleanTAP [1] and MSPASS [9], the availability of even basic implementations for full first-order modal logics is very limited.

Connection calculi are a well-known basis to automate formal reasoning in classical first-order logic [4, 11]. In connection calculi the proof search is guided by connections, i.e. sets of the form \{P, ¬P\} that contain literals with the same predicate symbol but different polarities. This yields a goal-oriented and, therefore, more efficient proof search. Connections correspond to closed branches in the tableau framework [8] or axioms in the sequent calculus [7]. The matrix characterisation for classical logic can be seen as the underlying notion of connection calculi for classical logic. This characterization of logical validity can be extended to modal logic by using prefixes, which encode the additional non-permutabilities of rule applications in the modal sequent calculus [16, 17].

In this paper an implementation of a connection calculus for the first-order modal logics D, S4, S5, and T is presented. It is based on the leanCoP system for classical first-order logic extended by an algorithm for prefix unification. This paper is structured as follows. In Section 2 the use of prefixes is motivated and the basic definitions required for the matrix characterization for modal logic are introduced. Section 3 presents the algorithm for prefix unification and the connection calculus used for the proof search. Details about the implementation based on this modal connection calculus are given in Section 4. Section 5 presents performance results of the implementation on the QMLTP problem library, before Section 6 concludes with a short summary and some remarks on future work.
2 Preliminaries

The reader is assumed to be familiar with the syntax and semantics of first-order modal logic, see, e.g., [6]. In the following we consider the standard semantics with rigid term designation, i.e. terms denote the same object in every world, and terms are local, i.e. any ground term denotes an existing object in every world. In this paper the letters $P, Q$ are used to denote predicate symbols, $f$ denotes a function symbol, and $x, y, z$ are used to denote (term) variables. Terms are denoted by $s, t$ and are built from functions, constants, and variables. Atomic formulae, denoted by $A$, are built from predicate symbols and terms. The connectives $\neg, \land, \lor, \Rightarrow$ denote negation, conjunction, disjunction and implication, respectively. A (first-order modal) formula, denoted by $F, G, H$, consists of atomic formulae, the connectives, the modal operators $\square$ and $\Diamond$, and the existential and universal quantifiers $\forall x$ and $\exists x$. A literal, denoted by $L$, has the form $A$ or $\neg A$. The complement $\overline{L}$ of a literal $L$ is $A$ if $L$ is of the form $\neg A$, and $\neg A$ otherwise.

2.1 Using Prefixes

In order to adapt the matrix characterization of classical validity to modal logic, the notion of prefixes and a modal substitution are introduced.

While the standard connectives and quantifiers are interpreted as in classical logic, the semantics of the modal operators $\square$ and $\Diamond$ is defined with respect to a set of worlds. Then, $\square F$ or $\Diamond F$ are true in a world $w$, if $F$ is true in all worlds accessible from $w$ or some world accessible from $w$, respectively. The properties of the accessibility relation determine the particular modal logic. In the following the modal logics $D, T, S4, and S5$ are considered. Their accessibility relation is serial ($D$)\(^1\), reflexive ($T$), reflexive and transitive ($S4$), or an equivalence relation ($S5$), respectively.

The modal sequent calculus extends the classical sequent calculus by the modal rules $\square$-left, $\square$-right, $\Diamond$-left and $\Diamond$-right, which are used to introduce the modal operators $\square$ and $\Diamond$ into the antecedent (left side) or the succedent (right side) of the sequent, respectively.

**Definition 1** (Modal sequent calculus). The sequent calculus for the modal logics $D, T, and S4$ consists of the axioms and rules of the classical sequent calculus and the four additional inference rules shown in Figure 1.\(^2\) It is $\Gamma_{\square} := \{\square G | \square G \in \Gamma\}$, $\Delta_{\Diamond} := \{\Diamond G | \Diamond G \in \Delta\}$, $\Gamma_{(\square)} := \{G | \square G \in \Gamma\}$, and $\Delta_{(\Diamond)} := \{G | \Diamond G \in \Delta\}$.

\[
\begin{array}{c l l l}
\Gamma^+, F \vdash \Delta^+ & \quad \Gamma^* \vdash F, \Delta^* & \quad \Gamma^* \vdash F, \Delta^* & \quad \Gamma^* \vdash F, \Delta^* \\
\Gamma, \square F \vdash \Delta & \quad \square \text{-left} & \quad \Gamma^+ \vdash \square F, \Delta & \quad \square \text{-right} \\
\Gamma^+ \vdash F, \Delta^+ & \quad \Diamond \text{-left} & \quad \Gamma, \Diamond F \vdash \Delta & \quad \Diamond \text{-right} \\
\Gamma^+ \vdash \square F, \Delta & \quad \Diamond \text{-right} & \quad \Gamma^* \vdash F, \Delta^* & \quad \square \text{-right} \\
\end{array}
\]

<table>
<thead>
<tr>
<th>Logic</th>
<th>$\Gamma^+$</th>
<th>$\Delta^+$</th>
<th>$\Gamma^*$</th>
<th>$\Delta^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D$</td>
<td>$\Gamma_{(\square)}$</td>
<td>$\Delta_{(\Diamond)}$</td>
<td>$\Gamma_{(\square)}$</td>
<td>$\Delta_{(\Diamond)}$</td>
</tr>
<tr>
<td>$T$</td>
<td>$\Gamma$</td>
<td>$\Delta$</td>
<td>$\Gamma_{(\square)}$</td>
<td>$\Delta_{(\Diamond)}$</td>
</tr>
<tr>
<td>$S4$</td>
<td>$\Gamma$</td>
<td>$\Delta$</td>
<td>$\Gamma_{(\square)}$</td>
<td>$\Delta_{(\Diamond)}$</td>
</tr>
</tbody>
</table>

Figure 1: The four additional rules of the modal sequent calculus

For the proof search the modal rules are applied upwards, i.e., in an analytic way. In that case formulae in the sets $\Gamma^+, \Delta^+, \Gamma^+$ and $\Delta^+$ are deleted from the premise. To avoid deleting formulae that are required for a proof the application order of the modal rules need to be controlled.

\(^1\)A relation $R \subseteq W \times W$ is serial iff for all $w_1 \in W$ there is some $w_2 \in W$ with $(w_1, w_2) \in R$.

\(^2\)The modal sequent calculus captures the cumulative domain condition (see Section 2.2). There are no similar cut-free sequent calculi (that have the subformula property) for the logics with constant or varying domain or for the modal logic $S5$. 

Example 1 (Application of modal rules). Let $F_1$ be the formula $\Box P \Rightarrow \Box \Box P$. The left side of Figure 2 shows a sequent derivation of $F_1$ in the modal sequent calculus for $S4$. When the rule $\Box$-left is applied first (in an analytic way), the following application of the rule $\Box$-right deletes the formula $P$ from the antecedent and the derivation cannot be turned into a proof. In the derivation of $F_1$ in the middle of Figure 2, the two rules $\Box$-right are applied before rule $\Box$-left, and the resulting derivation is a proof.

$$
\frac{\frac{\frac{P \vdash P}{\Box P \Rightarrow \Box \Box P}}{P \vdash \Box P}}{} \quad \frac{w_2 \vdash P}{w_2 \vdash \Box P} \quad \frac{\vdash \Box P \Rightarrow \Box \Box P}{\vdash P \Rightarrow \Box \Box P}
$$

$P : a_0 \vdash : a_0 a_1 a_2$ $\square$-left

$\Box : a_0 \vdash : a_0 a_1 a_2$ $\square$-right

$\Box : a_0 \vdash : a_0 a_1$ $\square$-right

$\Box : a_0 \vdash : a_0 a_1 a_2$ $\square$-right

$\vdash \Box P \Rightarrow \Box \Box P : a_0$

Figure 2: Sequent derivations of $\Box P \Rightarrow \Box \Box P$ for the modal logic $S4$

Semantically, every application of the rule $\square$-right introduces a new world accessible from the previous one. Starting with $w_0$, the derivation in the middle of Figure 2 is labelled with the worlds in which the sequent is forced. $\Box P$ is forced in $w_0$, $w_1$ is accessible from $w_0$ and $w_2$ is accessible from $w_1$. As the accessibility relation of $S4$ is transitive, $w_2$ is also accessible from $w_0$. Hence, the formula $\Box P$ on the left side of the sequent is preserved. For the modal logic $D$, whose accessibility relation is neither reflexive nor transitive, the applications of the rules $\square$-left (or $\square$-right) and $\square$-right would yield the upper sequent $\vdash P$. The $\Box P$ on the left side of the sequent does not survive more than one world transition as the accessibility relation for the modal logic $D$ is not transitive. In the following the notion of prefixes is used in order to name sequences of accessible worlds.

Definition 2 (Prefix). A prefix, denoted by $p$ or $q$, is a string (sequence of characters) over an alphabet $\mathcal{V} \cup \Pi$, in which $\mathcal{V}$ is a set of prefix variables and $\Pi$ is a set of prefix constants.

To distinguish between elements of $\mathcal{V}$ and elements of $\Pi$ within a prefix, elements of $\mathcal{V}$ are written in capitals, denoted by $V$ or $V_i$, whereas small letters, denoted by $a$ or $a_i$, are used to represent elements of $\Pi$. $u$, $w$ denote strings, $\varepsilon$ denotes the empty string, $u \circ w$ denotes the concatenation of $u$ and $w$. $u \preceq w$ holds if, and only if, $u$ is an initial substring of $w$ or $u = w$.

Example 2 (Prefixes for worlds). In the right derivation of $F_1$ in Figure 2, each formula is labelled with a prefix encoding the world in which the formula is forced. Starting with $a_0$, each application of a modal rule extends the prefix by one character. In the logic $S4$, $A_1$ could denote any world accessible from $a_0$. To ensure that the propositions $P$ on both sides of the sequent are forced in the same world, their two prefixes $a_0 A_1$ and $a_0 a_1 a_2$ need to refer to the same world. This is achieved by assigning the string $a_1 a_2$ to $A_1$. In the logic $D$, $A_1$ can only denote a world directly accessible from $a_0$. Hence, only the string $a_1$ can be assigned to $A_1$ and the two prefixes $a_0 A_1$ and $a_0 a_1 a_2$ cannot denote the same world.

Semantically, a prefix denotes a specific world in a model [5][17]. Prefixes of literals that form an axiom in the sequent calculus need to denote the same world. Proof-theoretically, a prefix of a formula $F$ captures the modal context of $F$ and specifies the sequence of modal rules that have to be applied (analytically) in order to obtain $F$ in the sequent. Prefix variables and constants represent applications of the modal rules $\square$-left/$\square$-right and $\square$-right/$\square$-left, respectively. When a modal rule that corresponds to an element in $\Pi$ (and $\mathcal{V}$ for the modal logic $D$) is applied, only formulae whose prefix ends with an element of $\mathcal{V}$ are preserved in the premise of the sequent. Hence, the application of the modal rules represented by elements of $\mathcal{V}$ and $\Pi$ need to be controlled. In order to preserve two atomic formulae that form an axiom in the sequent calculus their prefixes need to be made identical. This is done by a modal substitution that maps elements of $\mathcal{V}$ to strings over $\mathcal{V} \cup \Pi$.

20
Implementing Connection Calculi for First-order Modal Logics

**Definition 3** (Modal substitution). A modal substitution is a mapping \( \sigma_M : V \rightarrow (V \cup \Pi)^* \) that assigns a string over the alphabet \( V \cup \Pi \) to every element in \( V \). For the modal logics \( D \) and \( T \) the accessibility condition \( |\sigma_M(V)| = 1 \) or \( |\sigma_M(V)| \leq 1 \) has to hold for all \( V \in V \), respectively. \( \sigma_M(p) \) denotes the string \( p \) in which each prefix variable \( V \) is replaced by its image \( \sigma_M(V) \).

Substitutions \( \sigma_M \) are assumed to be idempotent, i.e. \( \sigma_M \circ \sigma_M = \sigma_M \) holds. They are represented by the set \( \{ V \setminus p \mid \sigma_M(V) = p \text{ and } p \neq V \} \). The accessibility condition encodes the characteristics of the different modal rules for each logic. For the modal logic \( D \) a formula that corresponds to a prefix variable will only survive one application of the modal rules. Hence, only strings of length one can be assigned to prefix variables. For the modal logic \( T \) a formula that corresponds to a prefix variable will only survive one application of the modal rules. Hence, only strings of length one or \( \epsilon \) can be assigned to prefix variables. For the modal logic \( S4 \) a formula that corresponds to a prefix variable will survive any number of applications of modal rules. Hence, there is no restriction on strings that can be assigned to prefix variables.

**Example 3** (Modal substitution). Consider the prefixes \( a_0 A_1 \) and \( a_0 a_1 a_2 \) of the two literals occurring in the upper sequent of the right derivation in Figure 2. \( \sigma_M = \{ A_1 \setminus a_1 a_2 \} \) is a modal substitution for the logic \( S4 \) (but not for \( D \) or \( T \)) with \( \sigma_M(a_0 A_1) = a_0 a_1 a_2 = g\sigma_M(a_0 a_1 a_2) \).

### 2.2 Matrix Characterization

The matrix characterization of validity for modal logic can be seen as the basis of the connection calculus presented in Section 3.2. As the core implementation of the connection calculus (introduced in Section 4.2) uses a clausal form, the original (non-clausal) matrix characterization [17] is slightly modified. Matrices use a clausal form and the definitions of prefixes and paths are simplified by using the formal definition of a prefixed matrix and an extended Skolemization technique.

A polarity \( pol \in \{ 0, 1 \} \) is used to represent negation in a matrix, i.e. literals of the form \( A \) and \( \neg A \) are represented by \( A^0 \) and \( A^1 \), respectively. A prefixed formula \( F^{pol} ; p \) is a formula \( F \) marked with a polarity \( pol \) and a prefix \( p \). A (prefixed) clause, denoted by \( C \), is a set \( \{ L_1 : p_1, \ldots, L_n : p_n \} \) with literals \( L_i \) and prefixes \( p_i \). \( F[x \lessdot t] \) denotes the formula \( F \) in which all free occurrences of \( x \) are replaced by \( t \).

**Definition 4** (Prefixed matrix). The prefixed matrix \( M(F^{pol} ; p) \) of a prefixed formula \( F^{pol} ; p \) is a set of prefixed clauses. It is defined inductively according to Table 7. The following abbreviation is used: \( M_G \cup \exists \beta M_H := \{ C_G \cup C_H \mid C_G \in G, C_H \in H \} \). \( x^* \) is a new term variable, \( t^* \) is the Skolem term \( f^*(x_1, \ldots, x_n) \) in which \( f^* \) is a new function symbol and \( x_1, \ldots, x_n \) are all free term and prefix variables in \( \forall x G \) or \( \exists x G \). \( V^* \) is a new prefix variable, \( a^* \) is a prefix constant of the form \( f^*(x_1, \ldots, x_n) \) in which \( f^* \) is a new function symbol and \( x_1, \ldots, x_n \) are all free term and prefix variables in \( \Box G \) or \( \Diamond G \).

The prefix of \( x^* \) and \( t^* \), denoted \( \text{pre}(x^*) \) and \( \text{pre}(t^*) \), is \( p \). The (prefixed) matrix \( M(F) \) of a (first-order modal) formula \( F \) is the prefixed matrix \( M(F^0 ; \varepsilon) \).

In the graphical representation of a matrix, its clauses are arranged horizontally, while the literals of each clause are arranged vertically. In order to simplify the implementation, Skolemization is not only used for Eigenvariables but also for prefix constants. A similar approach is already used for intuitionistic logic [10]. Thus, the irreflexivity test of the reduction ordering [17] is realized by the occur check of the term and prefix unification procedures. The same Skolem function symbol is used for instances of the same subformula, a technique that is similar to the liberalized \( \delta^+ \)-rule in classical tableau calculi [8].

**Example 4** (Prefixed matrix). Let \( F_1 \) be the modal formula \( (\Box \exists x P x \land \Box \forall y (\Diamond P y \Rightarrow Q y)) \Rightarrow \Box \exists z Q z \). The prefixed matrix of \( F_1 \) is \( M_1 := M(F_1) = \{ F^0 ; d : a_1 \}, \{ F^0 ; y : V_1 V_2, Q^0 ; y : V_3 \}, \{ Q^0 ; z : V_5 \} \). \( x \) is an Eigenvariable, \( y, z \) are free term variables, \( a_1 \in \Pi \) is a prefix constant, \( V_1, V_2, V_3 \in V \) are prefix variables. \( d, a_1 \) are Skolem functions/constants. The graphical representation of \( M_1 \) is shown in Figure 3.
Connections and paths are the basic concepts of the matrix characterization. They correspond to axioms and atomic sequents (that cannot be reduced anymore) in the sequent calculus, respectively.

**Definition 5** (Connection, path, term substitution). The set \( \{P^1(s_1, \ldots, s_n) : p_1, P^0(t_1, \ldots, t_n) : p_2\} \) is a connection. It consists of a pair of literals with the same predicate symbol but different polarities.

A path \( pt \) through a matrix \( M = \{C_1, \ldots, C_n\} \) is a set of literals that contains one literal from each clause, i.e. \( pt := \cup_{i=1}^n \{L_i\} \) with \( L_i \in C_i \). A term substitution \( \sigma_Q \) is a mapping from the set of term variables to the set of terms. In \( \sigma_Q(L) \) all term variables \( x \in L \) are substituted by their image \( \sigma_Q(x) \).

**Example 5** (Connection, path, term substitution). Consider the matrix \( M_1 \) of Example 4 and its graphical representation in Figure 3. Then, \( \{P^1 d : a_1, P^0 y : V_1 V_2\} \) and \( \{Q^1 P^0 y : V_1, Q^0 P^0 z : V_3\} \) are connections. \( \{P^1 d : a_1, P^0 y : V_1 V_2, Q^0 P^0 z : V_3\} \) and \( \{P^1 d : a_1, Q^1 P^0 y : V_1, Q^0 P^0 z : V_3\} \) are (the only) paths through \( M_1 \). \( \sigma_Q \) with \( \sigma_Q(1) = d \) and \( \sigma_Q(2) = d \) is a term substitution.

To ensure that a connection represents an axiom in the sequent calculus, the atomic formulae in the connection have to be unified under a term substitution. Furthermore, their prefixes have to unify under a prefix substitution. This ensures that these formulae are not deleted by applications of modal rules.

**Definition 6** (Admissible substitution, complementarity). Let \( \mathcal{L} \in \{D, T, S4, S5\} \) be a modal logic, \( \mathcal{D} \in \{\text{constant}, \text{cumulative}, \text{varying}\} \) be a domain, \( \sigma_M \) be a modal substitution for \( \mathcal{L} \) and \( \sigma_Q \) be a term substitution. The (combined) substitution \( \sigma := (\sigma_Q, \sigma_M) \) is admissible iff the domain condition in Table 2 holds for all term variables \( x \) and \( \mathcal{D} \) for cumulative domain for all Skolem terms \( t \) in \( \sigma_Q(x) \), or \( \mathcal{D} \) for varying domain for all term variables/Skolem terms \( s \) in \( \sigma_Q(x) \). A connection \( \{L_1 : p_1, L_2 : p_2\} \) is \( \sigma \)-complementary for a substitution \( \sigma = (\sigma_Q, \sigma_M) \) iff \( \sigma_Q(\sigma_M(L_1 : p_1)) = \sigma_Q(\sigma_M(L_2 : p_2)) \). For the modal logic \( S5 \) only the last character of all prefixes, i.e. of \( \text{pre}(x) \), \( \text{pre}(s) \), \( p_1 \) and \( p_2 \), is considered.

---

**Table 1: Matrix \( M(F^p; p) \) of a first-order modal formula \( F^p; p \)**

<table>
<thead>
<tr>
<th>type</th>
<th>( F^p; p )</th>
<th>( M(F^p; p) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>atomic</td>
<td>( A^0; p )</td>
<td>( {A^0; p} )</td>
</tr>
<tr>
<td></td>
<td>( A^1; p )</td>
<td>( {A^1; p} )</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>( (\neg G)^0; p )</td>
<td>( M(G^1; p) )</td>
</tr>
<tr>
<td></td>
<td>( (\neg G)^1; p )</td>
<td>( M(G^0; p) )</td>
</tr>
<tr>
<td></td>
<td>( (G \land H)^1; p )</td>
<td>( M(G^1; p) \cup M(H^1; p) )</td>
</tr>
<tr>
<td></td>
<td>( (G \lor H)^0; p )</td>
<td>( M(G^0; p) \cup M(H^0; p) )</td>
</tr>
<tr>
<td></td>
<td>( (G \Rightarrow H)^0; p )</td>
<td>( M(G^1; p) \cup M(H^0; p) )</td>
</tr>
<tr>
<td>( \nu )</td>
<td>( (\Box G)^1; p )</td>
<td>( M(G^1; p) \cup M^*(p) )</td>
</tr>
<tr>
<td></td>
<td>( (\Diamond G)^0; p )</td>
<td>( M(G^0; p) \cup M^*(p) )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>type</th>
<th>( F^p; p )</th>
<th>( M(F^p; p) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>( (G \land H)^0; p )</td>
<td>( M(G^0; p) \cup M(H^1; p) )</td>
</tr>
<tr>
<td></td>
<td>( (G \lor H)^1; p )</td>
<td>( M(G^1; p) \cup M(H^1; p) )</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>( (\forall x G)^1; p )</td>
<td>( M(G[x']^0; p) )</td>
</tr>
<tr>
<td></td>
<td>( (\exists x G)^0; p )</td>
<td>( M(G[x']^1; p) )</td>
</tr>
<tr>
<td>( \delta )</td>
<td>( (\forall x G)^0; p )</td>
<td>( M(G[x']^0; p) )</td>
</tr>
<tr>
<td></td>
<td>( (\exists x G)^1; p )</td>
<td>( M(G[x']^1; p) )</td>
</tr>
<tr>
<td>( \pi )</td>
<td>( (\Box G)^0; p )</td>
<td>( M(G^0; p \circ a^*) )</td>
</tr>
<tr>
<td></td>
<td>( (\Diamond G)^1; p )</td>
<td>( M(G^1; p \circ a^*) )</td>
</tr>
</tbody>
</table>

---

**Table 2: Domain condition for the substitution \( \sigma = (\sigma_Q, \sigma_M) \)**

<table>
<thead>
<tr>
<th>domain ( \mathcal{D} )</th>
<th>D, T, S4</th>
<th>S5</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>cumulative</td>
<td>( \sigma_M(\text{pre}(t)) \leq \sigma_M(\text{pre}(x)) )</td>
<td>–</td>
</tr>
<tr>
<td>varying</td>
<td>( \sigma_M(\text{pre}(s)) = \sigma_M(\text{pre}(x)) )</td>
<td>( \sigma_M(\text{pre}(s)) = \sigma_M(\text{pre}(x)) )</td>
</tr>
</tbody>
</table>
The constant domain variants place no restriction on the modal substitution as the same objects exist in every world. For the cumulative domain variants any Skolem term \( t \) introduced by a quantifier rule may only exists in the world in which they are introduced. Hence, if \( s \) occurs in \( \sigma_Q(x) \) then the prefixes \( w_1 = \sigma_M(\text{pre}(s)) \) and \( w_2 = \sigma_M(\text{pre}(x)) \) need to be identical.

Example 6 (Admissible substitution, complementarity). Consider the matrix \( M_1 \) of Example 4 with \( \sigma_Q(y) = \sigma_Q(z) = d, \sigma_M(V_1) = \sigma_M(V_3) = a_1, \) and \( \sigma_M(V_2) = \varepsilon \) (for \( T, S4 \)) or \( \sigma_M(V_2) = a_1 \) (for \( S5 \)). Then, \( \sigma_M \) is a modal substitution for the modal logics \( T, S4, \) and \( S5, \) but not for \( D. \) The substitution \( \sigma_1 = (\sigma_Q, \sigma_M) \) is admissible for the constant, cumulative, and varying domain as \( \text{pre}(y) = \text{pre}(z) = \text{pre}(d) = a_1. \) The connections \( \{P^1d : a_1, P^0y : V_1V_2\} \) and \( \{Q^1y : V_1, Q^0z : V_3\} \) are \( \sigma_1 \)-complementary.

The notion of multiplicity is used to encode the number of clause copies used in a proof. It is a function \( \mu : M \to \mathbb{N} \) that assigns a natural number to each clause in \( M \) specifying how many copies of this clause are considered in a proof. In the copy of a clause \( C \) every (term and prefix) variable in \( C \) is replaced by a unique new variable. \( M^\mu \) denotes the matrix that includes these clause copies. Clause copies correspond to applications of the contraction rule in the sequent calculus.

Theorem 1 (Matrix characterisation for modal logic). Let \( L \in \{D,T,S4,S5\} \) be a modal logic and \( D \in \{\text{constant},\text{cumulative},\text{varying}\} \) be a domain. A formula \( F \) is valid in the logic \( L \) with domain \( D \) iff there is a multiplicity \( \mu, \) an admissible substitution \( \sigma = (\sigma_Q, \sigma_M) \) for \( L/D, \) and a set of \( \sigma \)-complementary connections such that every path through \( M(F)^\mu \) contains a connection from this set.

Example 7 (Matrix characterisation for modal logic). Consider the formula \( F_1 \) of Example 2 and the graphical representation of its matrix \( M_1 \) in Figure 5. Let \( \mu = 1, \sigma_Q(y) = \sigma_Q(z) = d, \sigma_M(V_1) = \sigma_M(V_3) = a_1, \) and \( \sigma_M(V_2) = \varepsilon \) (for \( T,S4) \) or \( \sigma_M(V_2) = a_1 \) (for \( S5 \). \( \sigma_1 = (\sigma_Q, \sigma_M) \) is admissible for all domains, and every path through \( M_1^\mu \) contains a \( \sigma_1 \)-complementary connection. Hence, \( F_1 \) is valid in the modal logics \( T, S4, \) and \( S5 \) with constant, cumulative, and varying domain.

### 3 Proof Search

An effective way to carry out proof search that is based on the matrix characterization is done with a connection-driven search strategy. For this purpose the connection calculus is extended by additional prefixes. Furthermore, a prefix unification algorithm is required in order to calculate the modal substitution by unifying the prefixes of the literals in each connection.

#### 3.1 Prefix Unification

Formally, prefix unification is the problem to find a modal substitution or unifier \( \sigma_M \) for a set of prefix equations \( E = \{ p_1 = q_1, \ldots, p_n = q_n \} \) such that \( \sigma_M(p_i) = \sigma_M(q_i) \) for all \( 1 \leq i \leq n. \) A set of unifiers \( \Sigma \) is a set of most general unifiers for \( E \) if, and only if, every unifier \( \tau \) is an instance of some \( \sigma \in \Sigma \) (completeness) and no unifier \( \sigma \in \Sigma \) is an instance of another unifier \( \tau \in \Sigma \) (minimality).

---

3This condition — as well as the (non-applicable) accessibility condition for \( S5 \) — are slightly corrected conditions of \cite{i3}. 

---

Figure 3: Graphical representation of the matrix \( M_1 \)
As prefixes are strings, general algorithms for string unification could be used for unifying prefixes. But for general string unification the number of most general unifiers might not be finite. A more efficient unification takes the prefix property of all prefixes $p_1, p_2, \ldots$ into account: for all prefixes $p_i = u_i X w_1$ and $p_j = u_j X w_2$ with $X \in \mathcal{U} \cup \mathcal{I}$ the property $u_i = u_j$ holds. The prefix property reflects the fact that the order in which modal rules are applied corresponds to a (formula) tree structure.

Depending on the modal logic the accessibility condition (see Definition 3) has to be respected when calculating the modal substitution $\sigma_M$, i.e. for all $V \in \mathcal{U}$ : $|\sigma_M(V)| = 1$ for the modal logic $\mathcal{D}$, $|\sigma_M(V)| \leq 1$ for the modal logic $\mathcal{T}$ and no restriction for the modal logics $\mathcal{S}_4$ and $\mathcal{S}_5$. The prefix unification for $\mathcal{D}$ is a simple pattern matching, i.e. the standard term unification can be used. For $\mathcal{S}_4$ the prefix unification for intuitionistic logic can be used; see [10] for details. For $\mathcal{S}_5$ only the last character of each prefix (or $\varepsilon$ if the prefix is $\varepsilon$) has to be unified. The prefix unification for $\mathcal{T}$ is specified by a set of rewriting rules. Given a prefix equation $E = \{p = q\}$ it calculates a set of most general unifiers for $E$. For an equation of the form $V_1 u = V_2 w$ the substitution $\{V_1 \varepsilon, V_2 \varepsilon\}$ would not be minimal (but an instance of $\{V_1 \varepsilon, V_2 \varepsilon\}$). To prevent the calculation of such substitutions, the right side of each equation is split by a bar. Then, $V_2$ is moved to the left side of the bar, i.e., $u = V_2 | w$, when $\varepsilon$ is assigned to $V_1$.

**Definition 7** (Prefix unification). The prefix unification of two prefixes $p$ and $q$ for the modal logic $\mathcal{T}$ is done by applying the rewriting rules in Table 4 it is $V \in \mathcal{U}$, $a \in \mathcal{I}$, $X \in \mathcal{U} \cup \mathcal{I}$, $u, w \in (\mathcal{U} \cup \mathcal{I})^*$. The application of a rewriting rule replaces a tuple $(E, \sigma_M)$ by a tuple $(E', \sigma_M')$. $E$ and $E'$ are prefix equations, $\sigma_M$ and $\sigma_M'$ are modal substitutions. The unification starts with the tuple $(\{\}, \sigma_M)$ and terminates when the tuple $(\{\}, \sigma_M)$ is derived. In this case $\sigma_M$ represents a most general unifier. Rules can be applied non-deterministically as there might be more than one most general unifier.

<table>
<thead>
<tr>
<th>Rule</th>
<th>Premise</th>
<th>Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1.</td>
<td>$\varepsilon = \varepsilon</td>
<td>\varepsilon$, $\sigma_M$</td>
</tr>
<tr>
<td>R2.</td>
<td>$\varepsilon = \varepsilon</td>
<td>X w$, $\sigma_M$</td>
</tr>
<tr>
<td>R3.</td>
<td>$Vu = \varepsilon</td>
<td>\varepsilon$, $\sigma_M$</td>
</tr>
<tr>
<td>R4.</td>
<td>$Xu = \varepsilon</td>
<td>X w$, $\sigma_M$</td>
</tr>
<tr>
<td>R5.</td>
<td>$Vu = \varepsilon</td>
<td>X w$, $\sigma_M$</td>
</tr>
<tr>
<td>R6.</td>
<td>$au = \varepsilon</td>
<td>V w$, $\sigma_M$</td>
</tr>
<tr>
<td>R7.</td>
<td>$V_1 u = \varepsilon</td>
<td>V_2 w$, $\sigma_M$</td>
</tr>
<tr>
<td>R8.</td>
<td>$Vu = X</td>
<td>w$, $\sigma_M$</td>
</tr>
<tr>
<td>R9.</td>
<td>$Vu = X</td>
<td>w$, $\sigma_M$</td>
</tr>
<tr>
<td>R10.</td>
<td>$au = V</td>
<td>w$, $\sigma_M$</td>
</tr>
</tbody>
</table>

Figure 4: Prefix unification for the modal logic $\mathcal{T}$

To solve a set of prefix equations $E = \{p_1 = p_1, \ldots, q_n = t_q\}$ the equations in $E$ are solved one after the other. The modal substitution $\sigma_M$ calculated for the first equation is applied to $E$ and the next equation in $E$ is considered and so on. Different substitutions $\sigma_M$ are calculated on backtracking.

**Example 8** (Prefix unification). Consider the connection $\{P^1 : a_1, P^0 : V_1 V_2\}$ from Example 5 and 6.

The most general unifiers for its two prefixes $a_1$ and $V_1 V_2$ are calculated as follows: $\{a_1 = \varepsilon | V_1 V_2\}$, $\{\}$ $\xrightarrow{R5} \{V_1 V_2 = a_1 | \varepsilon\}$, $\{\}$ $\xrightarrow{R5} \{V_2 = a_1 | \varepsilon\}$, $\{V_1 \varepsilon\}$ $\xrightarrow{R5} \{\varepsilon = \varepsilon | \varepsilon\}$, $\{V_1 \varepsilon, V_2 \varepsilon\}$ $\xrightarrow{R1} \{\}, \{V_1 \varepsilon, V_2 \varepsilon\}$ and $\{a_1 = \varepsilon | V_1 V_2\}$, $\{\}$ $\xrightarrow{R5} \{V_1 V_2 = a_1 | \varepsilon\}$, $\{\}$ $\xrightarrow{R9} \{V_2 = \varepsilon | \varepsilon\}$, $\{V_1 \varepsilon, a_1\}$ $\xrightarrow{R3} \{\varepsilon = \varepsilon | \varepsilon\}$, $\{V_1 \varepsilon, V_2 \varepsilon\}$ $\xrightarrow{R1} \{\}, \{V_1 \varepsilon, V_2 \varepsilon\}$.

Hence, the two most general unifiers are $\{V_1 \varepsilon, V_2 \varepsilon\}$ and $\{V_1 a_1, V_2 \varepsilon\}$.

---

This is also called the monoid problem; it is the equation theory in which there is a neutral element $\varepsilon$ and the associativity of the string concatenation operator $\circ$ holds.

---

24
3.2 The Connection Calculus

The matrix characterization of Theorem 1 can serve as the basis for a proof calculus. Checking that all paths contain a \( \sigma \)-complementary connection can be done by, e.g., sequent or tableau calculi. The connection calculus uses a connection-driven search strategy in order to determine an appropriate set of connections. It is already successfully used for automated theorem proving in classical and intuitionistic logic \[11, 12\].

Proof search in the connection calculus starts by selecting a start clause. Afterwards, connections are successively identified in order to make sure that all paths through the matrix contain a \( \sigma \)-complementary connection for some combined substitution \( \sigma = (\sigma_Q, \sigma_M) \). This process is guided by an active path, a subset of a path through the matrix. The multiplicity \( \mu \) is increased dynamically during the path checking process. The connection calculus is adapted to modal logic by simply adding prefixes to all literals and using the prefix unification algorithm presented in Section 3.1.

Definition 8 (Modal connection calculus). The axiom and the rules of the modal connection calculus are given in Figure 5. The words of the calculus are tuples of the form “\( C, M, \text{Path} \)”, where \( M \) is a (prefixed) matrix, \( C \) and \( \text{Path} \) are sets of (prefixed) literals or \( \varepsilon \). \( C \) is called the subgoal clause and \( \text{Path} \) is called the active path. \( C_1 \) and \( C_2 \) are clauses, \( \sigma = (\sigma_Q, \sigma_M) \) is an admissible substitution, \( \{L_1 : p_1, L_2 : p_2\} \) is a \( \sigma \)-complementary connection. The substitutions \( \sigma_Q \) and \( \sigma_M \) are global (or rigid), i.e. they are applied to the whole derivation.

A derivation for \( C, M, \text{Path} \) with the admissible substitution \( \sigma = (\sigma_Q, \sigma_M) \) for the logic \( \mathcal{L} \) and the domain \( D \) in which all leaves are axioms is called a modal connection proof for \( C, M, \text{Path} \). A modal connection proof for the matrix \( M \) is a modal connection proof for \( \varepsilon, M, \varepsilon \).

Theorem 2 (Correctness and completeness). Let \( \mathcal{L} \in \{D, T, S4, S5\} \) be a modal logic and \( D \in \{\text{constant, cumulative, varying}\} \) be a domain. A modal first-order formula \( F \) is valid in the modal logic \( \mathcal{L} \) with domain \( D \) iff there is a modal connection proof for \( M(F) \) for the logic/domain \( \mathcal{L}/D \).

The proof of Theorem 2 is based on the matrix characterization for modal logic (see Theorem 1) and the correctness and completeness of the connection calculus \[4\]. Proof search in the clausal connection calculus is carried out by applying the rules of the calculus in an analytic way, i.e. from bottom to top. During the proof search backtracking might be required, i.e. alternative rules need to be considered if the chosen rule (instance) does not lead to a proof. Alternative applications of rules occur whenever
more than one rule or more than one instance of a rule can be applied, i.e. when choosing the clause $C_1$ in the start and extension rule or the literal $L_2 : p_2$ in the reduction and extension rule. No backtracking is required when choosing the literal $L_1 : p_1$ in the reduction or extension rule as all literals in $C$ are considered in subsequent proof steps anyway. The term substitution $\sigma_Q$ and the modal substitution $\sigma_M$ are calculated by algorithms for term unification and prefix unification (see Section 3.1), respectively, whenever a reduction or extension rule is applied.

Example 9 (Modal connection calculus). Consider the formula $F_1$ and its matrix $M_1 = \{\{P^1d : a_1\}, \{P^0y : V_1V_2, Q^1y : V_1\}, \{Q^0z : V_3\}\}$ of Example 4. A derivation for $M_1$ in the modal connection calculus with $\sigma_Q(z') = d$, $\sigma_M(V'_1) = \sigma_M(V'_3) = a_1$ and $\sigma_M(V'_2) = \varepsilon$ (for T, S4) or $\sigma_M(V'_2) = a_1$ (for S5) is shown in Figure 6. As all leaves are axioms and the substitution $\sigma_1 = (\sigma_Q, \sigma_M)$ is admissible the derivation is also a proof for $M_1$. Hence, the

\[
\begin{align*}
\frac{\{\}, M_1, \{P^1d : a_1, Q^1y' : V'_1\}}{A} & \quad \frac{\{\}, M_1, \{P^1d : a_1\}}{A} & \quad \frac{\{\}, M_1, \{P^1d : a_1\}}{A} \\
\frac{\{Q^1y' : V'_1\}, \{P^1d : a_1\}, \{P^0y : V_1V_2, Q^1y : V_1\}, \{Q^0z : V_3\}\} \quad \frac{\{\}, M_1, \{}, \{P^1d : a_1\}}{E} & \quad \frac{\{\}, M_1, \{\}, \{P^1d : a_1\}}{E} & \quad \frac{\{\}, M_1, \{\}, \{P^1d : a_1\}}{E} \\
\frac{\{P^1d : a_1\}}{\varepsilon} & \quad \frac{\{P^1d : a_1\}}{\varepsilon} & \quad \frac{\{\}, M_1, \{\}, \{P^1d : a_1\}}{S} \\
\varepsilon, \{\{P^1d : a_1\}, \{P^0y : V_1V_2, Q^1y : V_1\}, \{Q^0z : V_3\}\} & \quad \varepsilon, \{\{P^1d : a_1\}, \{P^0y : V_1V_2, Q^1y : V_1\}, \{Q^0z : V_3\}\} & \quad \varepsilon
\end{align*}
\]

Figure 6: A proof for $M_1$ in the modal connection calculus

The proof of $F_1$ is valid in the modal logics T, S4, and S5 (with constant, cumulative, and varying domain). The proof of $F_1$ using the graphical matrix representation is depicted in Figure 7. The literals of each connection are connected with a line. The literals of the active path are boxed. In the first step the first clause $\{P^1d : a_1\}$ is selected as start clause. The second step is an extension step to the first literal of the second clause. The second step is an extension step to the third clause. Variables in clauses that are used in a proof step are renamed.

Figure 7: The modal connection proof for $M_1$ using the graphical matrix representation

Additional techniques to prune the search space in the connection calculus are regularity, lemmata, restricted backtracking, and the use of a definition clausal form translation; see [12] for details. All these techniques can be integrated into the modal connection calculus as well.

4 An Implementation

The connection calculus for modal logic presented in Section 3.2 was implemented in Prolog. The main components are the connection-driven proof search procedure and the prefix unification algorithm. The complete source code can be obtained at http://www.leanco.de/mleanco/.
4.1 Prefix Unification

In the Prolog implementation strings are represented by Prolog lists, e.g. the prefix \(a_0a_1V_2\) is represented by the list \([a_0,a_1,V_2]\). Prefix variables are Prolog variables, i.e. they start with a capital, prefix constants are Prolog atoms, i.e. they start with a small letter.

The source code of the prefix unification for the modal logic T is shown in line R1 to R10 of Figure 8. Each Prolog clause corresponds exactly to one of the rewrite rules defined in Figure 4. The predicate \(\text{tuni}_\text{t}(S,[],T)\) succeeds if the two prefixes \(S\) and \(T\) can be unified with respect to the accessibility condition for the modal logic T given in Definition 3. In this case the prefix variables are instantiated with a most general unifier for \(S\) and \(T\). Alternative unifiers are calculated via backtracking. According to the accessibility condition for the modal logic T, prefix variables may only be instantiated with the empty string or a single prefix constant or variable. As the Skolemization technique is applied to prefix constants as well, a term unification with the Prolog predicate \(\text{unify}_\text{with}_\text{occurs}_\text{check}\) is required (line R4, R9 and R10) whenever a prefix constant needs to be unified with another prefix constant or variable.

```
(a) \text{prefix}_\text{unify}([]).
(b) \text{prefix}_\text{unify}([S=T|G]) :-
   \text{prefix}_\text{unify}(G).
(c) \{ -S2=S \rightarrow T2=T ; -S2=T , T2=S \}, \text{flatten}(S2,S1), \text{flatten}(T2,T1),
   \text{logic}(s5) \rightarrow \text{tuni}_\text{s5}(S1,T1);
   \text{logic}(d) \rightarrow \text{tuni}_\text{d}(S1,T1);
   \text{logic}(t) \rightarrow \text{tuni}_\text{t}(S1,[],T1);
   \text{logic}(s4) \rightarrow \text{tunify}(S1,[],T1),
   \text{prefix}_\text{unify}(G).
(R1) \text{tuni}_\text{t}([],[],[]).
(R2) \text{tuni}_\text{t}([],[],[X|T]) :- \text{tuni}_\text{t}([X|T],[],[]).
(R3) \text{tuni}_\text{t}([V|S],[],[]) :- V=[], \text{tuni}_\text{t}(S,[],[]).
(R4) \text{tuni}_\text{t}([X1|S],[],[X2|T]) :- (\text{var}(X1) \rightarrow (\text{var}(X2), X1=X2);
   (+\text{var}(X2), \text{unify}_\text{with}_\text{occurs}_\text{check}(X1,X2))),!
   \text{tuni}_\text{t}(S,[],T).
(R5) \text{tuni}_\text{t}([V|S],[],[X|T]) :- \text{var}(V), \text{tuni}_\text{t}([V|S],[X],T).
(R6) \text{tuni}_\text{t}([C|S],[],[V|T]) :- \{ +\text{var}(C), \text{var}(V), \text{tuni}_\text{t}([V|T],[C],S) \}.
(R7) \text{tuni}_\text{t}([V1|S],[],[V2|T]) :- \text{var}(V1), V2=[], \text{tuni}_\text{t}([V1],S).
(R8) \text{tuni}_\text{t}([V|S],[X],T) :- V=[], \text{tuni}_\text{t}(S,[X],T).
(R9) \text{tuni}_\text{t}([V|S],[X],T) :- \text{var}(V), \text{unify}_\text{with}_\text{occurs}_\text{check}(V,X),
   \text{tuni}_\text{t}(S,[],T).
(R10) \text{tuni}_\text{t}([C|S],[V],T) :- \{ +\text{var}(C), \text{var}(V), \text{unify}_\text{with}_\text{occurs}_\text{check}(V,C),
   \text{tuni}_\text{t}(S,[],T).
```

Figure 8: Source code of the prefix unification for the modal logic T

The predicate \(\text{prefix}_\text{unify}(G)\) is used to solve a set of prefix equations. Its code is shown in the upper part of Figure 8. The set of prefix equations \(G\) is presented by a Prolog list of equations of the form \(S=T\). In each step one prefix equation is selected (line b) and the appropriate prefix unification is invoked depending on the logic specified by the \(\text{logic}\) predicate (line d to g). The remaining equations are considered afterwards (line h). If all equations are solved (line a) the predicate \(\text{prefix}_\text{unify}\) succeeds. A minus sign in front of a prefix (used for technical reasons) is deleted and prefixes are flattened (as they may contain nested lists) before the unification starts (line c).

The predicates \(\text{tuni}_\text{s5}\) and \(\text{tuni}_\text{d}\) implementing the prefix unification for the modal logics S5 and D, respectively, are straightforward. Details about the predicate \(\text{tunify}\) implementing the prefix unification for the modal logic S4 can be found in [10].
4.2 The Connection Calculus

The implementation of the connection calculus for first-order modal logic, called MleanCoP, is based on leanCoP 2.0, an automated theorem prover for first-order classical logic \cite{11}. To adapt the implementation to the modal connection calculus of Section 3.2, the leanCoP prover is extended by (a) prefixes that are added to literals and collected during the proof search and (b) an additional list that contains term variables together with their prefixes in order to check the domain condition. First, MleanCoP performs a classical proof search. After a proof is found, the prefixes of the literals in each connection are unified and the domain condition is checked. If the prefix unification fails the (classical) search for alternative connections continues via backtracking. The Prolog source code of the MleanCoP core prover is shown in Figure 9. The underlined code was added to leanCoP; no other modifications were done. In ECLiPSe Prolog sound unification has to be switched on with set\_flag(occur\_check, on).

```
prove(PathLim,Set) :-
  ( \+member(scut,Set) ->
    prove([(-#):[\-[]]],[],PathLim,[],[PreSet,FreeV1],Set) ;

dlit((#):_FV:_,_) ->
  prove(C,[(-#):[\-[]]],PathLim,[],[PreSet,FreeV1],Set),
  append(FreeV,FV,FreeV1).
 }

prove(PathLim,Set) :-
  member(comp(Limit),Set), PathLim\=Limit -> prove(1,[]);
  (member(comp(_),Set);retract(pathlim)) ->
  PathLim\=PathLim+1, prove(PathLim,Set).

prove([Lit:Pre|Cla],Path,PathLim,Lem,[PreSet,FreeV]) :-
  \+ (member(LitC,[Lit:Pre|Cla]), member(LitP,Path), LitC==LitP),
  \+ (NegLit=Lit;Lit=NegLit) ->
  ( member(LitL,Lem), Lit:Pre==LitL, PreSet3=[], FreeV3=[] ;
    member(NegLit:PreN,Path), unify_with_occurs_check(NegLit,NegLit),
    \+ prefix_unify([Pre=PreN]), PreSet3=[Pre=PreN], FreeV3=[] ;
    lit(NegLit:PreN,FV:Cl,Gnd1),
    \+ \+ prefix_unify([Pre=PreN]),
    \+ (Gnd1\=g -> true ; length(Path,K), K<PathLim \= true ;
      \+ pathlim \= assert(pathlim), fail ),
    prove(Cla,[Lit:Pre|Path],PathLim,Lem,[PreSet,FreeV],Set),
    PreSet3=[Pre=PreN|PreSet1], append(FreeV,FV,FreeV3) )
  ),

prove(PathLim,Lem,[Lit:Pre|Lem],Set) :-
  member(cut,Set) -> ! ; true,
  prove(Cla,[Lit:Pre|Path],Lem,[PreSet2,FreeV2],Set),
  append(PreSet3,PreSet2,PreSet), append(FreeV2,FreeV3,FreeV).
```

Figure 9: Source code of the MleanCoP core prover

The open subgoal $C$ and the active path $Path$ in the modal connection calculus of Figure 5 are represented by the Prolog lists $Cla$ and $Path$, respectively. The matrix $M$ is written into Prolog’s database in a preprocessing step. For every clause $C\in M$ and for every literal $lit(Cla,Path,Lem,[PreSet,FreeV])$ is stored where $FV$ is a list of elements of the form $[V,Pre]$ in which $V$ is a term variable in $C$ and $Pre$ is its prefix, $Cla=C\setminus \{Lit:Pre\}$ and $Gnd$ is $g$ if $C$ is ground, otherwise $Gnd$ is $n$. Atoms are represented by Prolog atoms, term variables by Prolog variables and negation by "\-". The substitutions $\sigma_Q$ and $\sigma_M$ are stored implicitly by Prolog.
The predicate `prove(Cla, Path, PathLim, Lem, [PreSet, FreeV], Set)` in line 1 to 19 implements the axiom, the reduction rule and the extension rule of the modal connection calculus of Figure 5. `PathLim` is the maximum size of the active path used for iterative deepening. `Lem` is a set of literals used as lemmata, and `Set` is a list of options used to control the proof search (see for details). For modal logic the lists `PreSet` and `FreeV` were added. They contain a list of prefix equations and a list of term variables with their prefixes that are collected during the proof search.

Line 1 implements the axiom, line 4 calculates the complement of the first literal `Lit:Pre` in the list of open subgoals, which is used for the following reduction step (line 7/8 and 18/19) or extension step (line 10/11, 14/15 and 18/19). In line 7 it is checked whether the active path `Path` contains a literal `NegLit:Pre` that unifies with the complement `NegLit:Pre` of the literal `Lit:Pre`. In line 8 a weak prefix unification of the prefix equation `Pre=Pre` is carried out (double negation prevents any variable bindings). Lines 18 and 19 continue the proof search for the premise of the reduction rule.

In line 10 the predicate `lit(NegLit:PreN,FV:Clai,Grnd1)` is used to find a clause that contains the complement of the literal `Lit:Pre`. `FV` is the list of term variables occurring in `Clai`. `Clai` is the remaining set of literals of the selected clause. Again, weak prefix unification in line 11 ensures that the prefixes of the new connection are unifiable. In line 14 the proof search for the left premise of the extension rule is invoked with the new open subgoal `Clai` and the active path `Path` extended by the literal `Lit:Pre`. Afterwards the proof search for the right premise is invoked in line 18. The lists of prefix equations and term variables are collected in line 8, 15 and 19.

The predicate `prove(PathLim,Set)` in line a to k implements the start rule of the modal connection calculus in Figure 5. When the matrix `M` is written into Prolog’s database the distinct literal `#[]` is added to all positive clauses. In line `c` the proof search starts with the open subgoal `[((-())]:([-[]])` and the empty active path `[ ]`, i.e. only positive clauses are used as start clause. After the classical proof search succeeds the domain condition is checked and the prefixes of each connection are unified. To this end the two predicates `domain_cond` (see the complete source code on the `MleanCoP` web site) and `prefix_unify` (see Section 4.1) are invoked in line `g`.

Prolog uses a depth-first search strategy resulting in an incomplete proof search. In order to regain completeness `MleanCoP` performs an iterative deepening on the size of the active path. When the extension rule is applied and the new clause is not ground, i.e. it contains a (term or prefix) variable, it is checked whether the size `K` of the active path exceeds the current path limit `PathLim` (line 12). In this case the predicate `pathlim` is written into Prolog’s database (line 13) indicating the need to increase the path limit. If the proof search for the current path limit fails and the predicate `pathlim` was written into the database (line `j`), then `PathLim` is increased and the proof search starts again (line `k`).

`MleanCoP` uses a few additional techniques, which are integrated into `MleanCoP` as well; see for details. `Regularity` ensures that no literal occurs more than once in the active path. It is integrated into the modal connection calculus in Figure 5 by imposing the following restriction on the reduction and the extension rule: `∀ L′:p′ ∈ C∪{L; p₁} : σQ(σM(L′; p′)) /∈ σQ(σM(Path))`. Line 3 implements a ground version of the regularity condition in `MleanCoP`. The idea of `Lemmata` or factorization is to reuse subproofs during the proof search. To this end an additional argument `Lem` is added to the `prove` predicate and line 5 is added to the `MleanCoP` source code. `Restricted backtracking` is a very effective technique to restrict backtracking in the connection calculus. It is implemented in line `b`, `d`, `e`, `f`, and `17`. A definition `clausal form` translation reduces the number of possible connections and, hence, the search space. Furthermore, `MleanCoP` uses a fixed strategy scheduling, i.e. the core prover is invoked repeatedly by a shell script with different prover options given in the list `Set`.

Example 10 (MleanCoP). Consider the modal formula `F₁` and its matrix `M₁` of Example 4. The goal `prove(((: ex X:p(X) , #: all Y; ( p(Y) -> q(Y) ) ) -> : ex Z:q(Z)))` succeeds; hence, `F₁` is valid. The (internal) representation of `M₁` (with added variable/prefix lists) is `[[1]:(-(p(d"a1")):a1)): -(a1)), ((Y,[V1]]): [p(Y): [V1,V2], -(q(Y)): -(V1)), ([Z,[V3]]): [q(Z): [V3]]].`
5 Performance

The MleanCoP implementation of Section 4.2 was tested on all 580 uni-modal problems of version 1.1 of the QMLTP library [14]. All problems were converted into the MleanCoP syntax using the TPTP2X tool [15] together with the format file included in the QMLTP library. All tests were conducted on a 3.4 GHz Xeon system with 4 GB of RAM running Linux 2.6.24 and ECLiPSe Prolog 5.10. The CPU time limit for all proof attempts was 600 seconds. Table 3 gives a summary of the test result. It shows the number of solved problems for the varying, cumulative, and constant domain of the modal logics D, T, S4, and S5 for the provers MleanCoP 1.2 (CoP), MleanTAP 1.3 (TAP), and MleanSeP 1.2 (SeP).

Table 3: Number of solved problems from the QMLTP library v1.1

<table>
<thead>
<tr>
<th>logic</th>
<th>varying domain</th>
<th>cumulative domain</th>
<th>constant domain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CoP</td>
<td>TAP</td>
<td>SeP</td>
</tr>
<tr>
<td>D</td>
<td>422</td>
<td>104</td>
<td>–</td>
</tr>
<tr>
<td>T</td>
<td>369</td>
<td>142</td>
<td>–</td>
</tr>
<tr>
<td>S4</td>
<td>393</td>
<td>173</td>
<td>–</td>
</tr>
<tr>
<td>S5</td>
<td>447</td>
<td>223</td>
<td>–</td>
</tr>
</tbody>
</table>

MleanSeP implements the standard modal sequent calculus (see Definition 1) for several modal logics. Proof search is carried out in an analytic way and free term variables are used together with a dynamic Skolemization to ensure the Eigenvariable condition. For the constant domain variants the Barcan formula (scheme) is added in a preprocessing step. The prover does not deal with the modal logic S5 and the varying domain variants as they do not have (cut-free) sequent calculi [17].

MleanTAP is a compact implementation of a prefixed tableau calculus for several modal logics. Similar to the modal connection calculus presented in Section 3.2 it is based on the modal matrix characterization of validity given in Theorem 1. Like MleanCoP it first performs a purely classical proof search. After a classical tableau proof is found the prefixes of those literals that closed the tableau branches are unified. The existence of a unifier ensures that the given formula is valid in modal logic. MleanCoP and MleanTAP use the same source code for the prefix unification and for checking the domain condition.

Table 4 presents detailed performance results of the three theorem provers on the QMLTP library. It shows for each considered modal logic and domain the number of solved problems, the number of proved (i.e. valid) problems, the number of refuted (i.e. invalid) problems, and the number of problems that are proved within specific time intervals. In general, MleanCoP proves and refutes significantly more problems than MleanTAP and MleanSeP. The number of proved problems by MleanTAP and MleanSeP is similar, but MleanSeP has a better time complexity behaviour; MleanTAP solves only two problems (for S4) in a time of more than 10 seconds. As MleanCoP uses a fixed strategy scheduling, many problems are solved in a time of more than one second. In general, the fewest problems are proved for the modal logic D with varying domain and the highest problems are proved for S5 with cumulative or constant domain. This behaviour is in line with the following fact:

If $\text{Theorem}(\mathcal{L}, D)$ is the set of theorems in the modal logic $\mathcal{L}$ with the domain $D$, then for all logics $\mathcal{L} \in \{D, T, S4\}$ and all domains $D$ it is: $\text{Theorem}(D, D) \subset \text{Theorem}(T, D) \subset \text{Theorem}(S4, D) \subset \text{Theorem}(S5, D)$ and $\text{Theorem}(\mathcal{L}, \text{varying}) \subset \text{Theorem}(\mathcal{L}, \text{cumulative}) \subset \text{Theorem}(\mathcal{L}, \text{constant})$; furthermore, it is $\text{Theorem}(S5, \text{varying}) \subset \text{Theorem}(S5, \text{cumulative}) = \text{Theorem}(S5, \text{constant})$.

---

5None of these provers were tuned towards the problems in the QMLTP library. MleanSeP and MleanTAP can be obtained at [http://www.leancop.de/mleansep/](http://www.leancop.de/mleansep/) and [http://www.leancop.de/mleantap/](http://www.leancop.de/mleantap/).
Table 4: Detailed performance results for the modal logics D, T, S4, and S5

<table>
<thead>
<tr>
<th>Logic</th>
<th>Varying domain</th>
<th>Cumulative domain</th>
<th>Constant domain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CoP</td>
<td>TAP</td>
<td>SeP</td>
</tr>
<tr>
<td>D</td>
<td>solved</td>
<td>422</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>proved</td>
<td>179</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>refuted</td>
<td>243</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>proved 0s to 1s</td>
<td>149</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>1s to 10s</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>10s to 100s</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>100s to 600s</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>T</td>
<td>solved</td>
<td>369</td>
<td>142</td>
</tr>
<tr>
<td></td>
<td>proved</td>
<td>224</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td>refuted</td>
<td>145</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>proved 0s to 1s</td>
<td>199</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td>1s to 10s</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>10s to 100s</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>100s to 600s</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>S4</td>
<td>solved</td>
<td>393</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>proved</td>
<td>274</td>
<td>169</td>
</tr>
<tr>
<td></td>
<td>refuted</td>
<td>119</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>proved 0s to 1s</td>
<td>245</td>
<td>167</td>
</tr>
<tr>
<td></td>
<td>1s to 10s</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>10s to 100s</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>100s to 600s</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>S5</td>
<td>solved</td>
<td>447</td>
<td>223</td>
</tr>
<tr>
<td></td>
<td>proved</td>
<td>359</td>
<td>219</td>
</tr>
<tr>
<td></td>
<td>refuted</td>
<td>88</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>proved 0s to 1s</td>
<td>307</td>
<td>215</td>
</tr>
<tr>
<td></td>
<td>1s to 10s</td>
<td>23</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>10s to 100s</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>100s to 600s</td>
<td>33</td>
<td>0</td>
</tr>
</tbody>
</table>

6 Conclusion

An implementation of a modal connection calculus for various first-order modal logics was presented. The modal connection calculus extends the classical calculus by prefixes and an additional prefix unification procedure. The specific modal logic and domain variant is selected through different prefix unification algorithms and different domain conditions, respectively. The Skolemization technique is extended to prefix variables. This makes an explicit reflexivity test of the reduction ordering redundant and simplifies the implementation. The implementation itself is based on the classical leanCoP prover. Optimization techniques used in leanCoP, such as regularity, lemmata, restricted backtracking and a definitional clausal form translation, are adapted and integrated into the modal prover MleanCoP as well.

Experimental results on the QMLTP library indicate that the performance of MleanCoP is significantly higher than the performance of provers based on analytic sequent calculi or prefixed tableau calculi. Hence, a connection-driven strategy seems to be essential for an efficient proof search based on
Approaches that use prefixes and a prefix unification procedure. A similar effect was already observed for first-order intuitionistic logic [10].

The development of automated theorem proving systems for first-order modal logic is still in its early stages. Only few other implementations exist so far. Another promising approach is the embedding of first-order modal logic into simple type theory [2] [3]. Future work includes the comparison with implementations of this and other approaches.

While a clausal form technically simplifies the proof search procedure, the translation into clausal form has a negative effect on the number of (possible) connections and, thus, on the size of the resulting search space. A non-clausal connection calculus [13] avoids any translation steps and preserves the structure of the original formula. Future work also includes the improvement of the prefix unification algorithm as this is the component that captures the main part of the additional search required for (first-order) modal logics.

Acknowledgements. The author would like to thank Thomas Raths for testing all implementations on the QMLTP library and providing the presented performance statistics.

References