Disjunction and Modular Goal-directed Proof Search[1]

Matthew Stone
Department of Computer Science and Center for Cognitive Science
Rutgers University
110 Frelinghuysen Road, Piscataway NJ 08854-8019
mdstone@cs.rutgers.edu

Summary

This paper explores goal-directed proof search in first-order multi-modal logic. The key issue is to design a proof system that respects the modularity and locality of assumptions of many modal logics. By forcing ambiguities to be considered independently, modular disjunctions in particular can be used to construct efficiently executable specifications in reasoning tasks involving partial information that otherwise might require prohibitive search. To achieve this behavior requires prior proof-theoretic justifications of logic programming to be extended, strengthened, and combined with proof-theoretic analyses of modal deduction in a novel way.

Contents

[1]Thanks to three anonymous reviewers, Mark Steedman, Rich Thomason, L. Thorne McCarty and Michael Fourman for extensive comments. This work was supported by an NSF graduate fellowship, an IRCS graduate fellowship, and a postdoctoral fellowship from RUCCS, as well as NSF grant IRI95-04372, ARPA grant N6601-94-C6043, and ARO grant DAAH05-94-G0426. June 4, 2002.

1 Introduction

This paper explores the proof-theoretic interaction between the goal-directed application of logical inferences and *information-flow*—that is, the possible connections between assumptions and conclusions in proofs. My own starting-point for this exploration was the result of [Stone, 1999], that intuitionistic sequent calculi can be formulated so as exhibit the characteristically intuitionistic *modular* information-flow (as underlying the correspondence between proofs and programs [Howard, 1980], for example) while nevertheless allowing logical inferences to be applied in any order whatsoever. This raises the question whether it is possible to enforce this kind of modularity *incrementally* during goal-directed proof search. Of course, the well-known flexibility of deduction in nonclassical logics [Fitting, 1972, Fitting, 1983, Wallen, 1990] is ample motivation for the question.

1.1 Problem Statement

I begin by delineating the focus of the paper more precisely. I will work with a family of first-order multi-modal logics in this paper. The generalization from intuitionistic logic reflects the utility of more general ways of structuring logical specifications [Baldoni et al., 1993, Baldoni et al., 1998a], as well as the broader importance of expressive modality in practical knowledge representation [McCarthy, 1993, McCarthy, 1997]. Qualitatively, what distinguishes the logics I consider (for which formal definitions are provided in Section 2) is that they permit rules of modal inference to be formulated in two equivalent ways [Fitting, 1972, Fitting, 1983, Wallen, 1990]. I illustrate the alternatives for the case of S4, a logic that we can perhaps regard as the pure modal logic of local and global modular assumptions [Giordano and Martelli, 1994].

1.1.1 Structural scope and modularity

The first formulation of modal inference is illustrated by the sequent inference figure below:

$$\frac{\Gamma^* \rightarrow G, \Delta^*}{\Gamma \rightarrow \Box G, \Delta}\,(\rightarrow \Box)$$

Such inferences set up a discipline of structural scope in proofs. Read upward, as a description of proof search, the figure describes how to accomplish generic reasoning about a modal context, such as the conclusion $\Box G$ here. We have to transform the sequent we are considering, by restricting our attention just to the generic modal statements in the sequent. Specifically, Γ^* contains the formula occurrences of the form $\Box A$ in Γ, and Δ^* contains the formula occurrences of the form $\Diamond A$ in Δ. The effect of the transformation is that we move from our current scope into a new, nested scope in which just generic information is available. Figure 1 illustrates all the structurally-scoped S4 sequent inferences that I will draw on in this motivating discussion; I refer the reader to [Fitting, 1983, Wallen, 1990] for more details on structurally-scoped proof.

The ability to define structural scope is intimately connected with the ability to describe modular and local reasoning. In specifying reasoning, we can think of antecedent formulas in sequents as program statements and succedent formulas in sequents as goals. In modal logics with structural scope, a necessary goal $\Box G$ can be seen as a *modular* goal because, as enforced by the structurally-scoped inference figure, only program statements of the form $\Box P$ can contribute to its proof. In other words, we cannot use the entire program to prove G; rather, we must use a designated *part* of the program: formulas of the form $\Box P$. This is the *module* we use to prove G. Multi-modal logic

$$\frac{\Gamma^* \rightarrow G, \Delta^*}{\Gamma \rightarrow \Box G, \Delta} (\rightarrow \Box)$$

$$\frac{\Gamma, G \rightarrow \Delta}{\Gamma, \Box G \rightarrow \Delta} (\Box \rightarrow)$$

$$\frac{\Gamma, P \rightarrow G, \Delta}{\Gamma \rightarrow P \supset G, \Delta} (\rightarrow\supset)$$

$$\frac{\Gamma \rightarrow G, \Delta \qquad \Gamma, P \rightarrow \Delta}{\Gamma, G \supset P \rightarrow \Delta} (\supset\rightarrow)$$

$$\frac{\Gamma, A \rightarrow \Delta \qquad \Gamma, B \rightarrow \Delta}{\Gamma, A \vee B \rightarrow \Delta} (\vee \rightarrow)$$

$$\frac{\Gamma \rightarrow A, \Delta \qquad \Gamma \rightarrow B, \Delta}{\Gamma \rightarrow A \wedge B, \Delta} (\rightarrow \wedge)$$

$$\Gamma, A \rightarrow A, \Delta \quad (Axiom)$$

Figure 1: Six inference figures and the axiom for structurally-scoped S4. After [Fitting, 1983, Wallen, 1990]. Sequents are multisets of modal formulas; this formulation (though not others that we will consider later) requires a structural rule of contraction. See Section 2.

allows us to name modules in a general way [Baldoni et al., 1993, Baldoni et al., 1998a].

In fragments of logic without the operator $\Diamond$, including S4 translations of intuitionistic formulas in particular, modularity brings *locality*. A goal $\Box(P \supset G)$ introduces a *local* assumption P. The assumption is local in the sense that it can only contribute to the proof of G, and cannot contribute to any other goal. We can motivate this locality in logical terms by examining the sequent inferences for $(\rightarrow \Box)$ and $(\rightarrow\supset)$ in combination:

$$\dfrac{\dfrac{\Gamma^*, P \longrightarrow G}{\Gamma^* \longrightarrow P \supset G} (\rightarrow\supset)}{\Gamma \longrightarrow \Box(P \supset G), \Delta} (\rightarrow \Box)$$

Observe that this logical fragment is constructed so that the succedent context Δ^* above $(\rightarrow \Box)$ is empty, and so we introduce P into a subproof where G is the only goal.

Logical modularity and locality underlie the use of the proof theory of modal logic as a declarative framework for structuring specifications, and thereby facilitating their design and reuse [Miller, 1989, Giordano and Martelli, 1994, Baldoni et al., 1993, Baldoni et al., 1996, Baldoni et al., 1998a].[2] Concretely, a goal that specifies the part of the program to be used in its proof will give rise to the same operational behavior when other parts of the program change. In this paper, I further emphasize that logical modularity and locality provide declarative tools for constraining the complexity of proof search itself. My motivating example is the proof in Figure 2,

[2]The model theory of modal logic can also be used to structure specifications [Sakakibara, 1987].

$$
\begin{prooftree}
\AxiomC{$A,\ldots \longrightarrow A$}
\AxiomC{$B,\ldots \longrightarrow B,A$}
\AxiomC{$A,\ldots \longrightarrow A$}
\RightLabel{$(\supset\rightarrow)$}
\BinaryInfC{$B,B\supset A,\ldots \longrightarrow A$}
\RightLabel{$(\Box\rightarrow)$}
\UnaryInfC{$B,\Box(B\supset A),\ldots \longrightarrow A$}
\RightLabel{$(\vee\rightarrow)$}
\BinaryInfC{$A\vee B,\Box(B\supset A),\ldots \longrightarrow A$}
\RightLabel{$(\Box\rightarrow)$}
\UnaryInfC{$\Box(A\vee B),\Box(B\supset A),\ldots \longrightarrow A$}
\RightLabel{$(\rightarrow\Box)$}
\UnaryInfC{$\Box(A\vee B),\Box(B\supset A),\ldots \longrightarrow \Box A$}
\AxiomC{$C,\ldots \longrightarrow C$}
\AxiomC{$D,\ldots \longrightarrow D,C$}
\AxiomC{$C,\ldots \longrightarrow C$}
\RightLabel{$(\supset\rightarrow)$}
\BinaryInfC{$D,D\supset C,\ldots \longrightarrow C$}
\RightLabel{$(\Box\rightarrow)$}
\UnaryInfC{$D,\Box(D\supset C),\ldots \longrightarrow C$}
\RightLabel{$(\vee\rightarrow)$}
\BinaryInfC{$C\vee D,\Box(D\supset C),\ldots \longrightarrow C$}
\RightLabel{$(\Box\rightarrow)$}
\UnaryInfC{$\Box(C\vee D),\Box(D\supset C),\ldots \longrightarrow C$}
\RightLabel{$(\rightarrow\Box)$}
\UnaryInfC{$\Box(C\vee D),\Box(D\supset C),\ldots \longrightarrow \Box C$}
\RightLabel{$(\rightarrow\wedge)$}
\BinaryInfC{$\Box(A\vee B),\Box(C\vee D),\Box(B\supset A),\Box(D\supset C) \longrightarrow (\Box A)\wedge(\Box C)$}
\end{prooftree}
$$

Figure 2: This structurally-scoped S4 proof shows how the locality of modular assumptions limits the possible interactions in proof. Ellipses mark points in sequents where I have suppressed additional formula occurrences that no longer contribute to the inference.

which establishes that the conclusion

$$(\Box A)\wedge(\Box C)$$

follows from the assumptions

$$\Box(A\vee B),\Box(C\vee D),\Box(B\supset A),\Box(D\supset C)$$

The assumptions in this proof—the program statements—specify two ambiguities. Either A or B holds, and either C or D holds. As part of the specification, we use modal operators to say how to reason with these ambiguities: we have $\Box(A\vee B)$ and $\Box(C\vee D)$. This means that the ambiguities themselves are *generic*; we can use them to perform case analysis at any time. However, when we reason about any particular case, we make *local* assumptions—we will assume A rather than $\Box A$ for example.

This specification limits the way case analyses in the proof interact. Consider our goal here: $(\Box A)\wedge(\Box C)$. We must prove each conjunct separately, *using generic information*; that is, each conjunct is proved in its own new nested scope. Thus, in the proof of Figure 2, we perform case analysis from $\Box(A\vee B)$ within the nested scope for $\Box A$, and perform case analysis from $\Box(C\vee D)$ within the nested scope for $\Box C$. Observe that the logic dictates the choice for us. For instance, performing case analysis from $\Box(A\vee B)$ within the nested scope for $\Box C$ is useless—the assumption of A and B cannot help here. Importantly, performing case analysis from $\Box(A\vee B)$ at the initial, outermost scope is also useless. Whatever assumptions we make will have to be discarded when we try to prove the conjuncts, and consider only generic information. This specification therefore cordons off the two ambiguities from one another in this proof problem. We have to consider the ambiguities separately.

Effectively, it is part of the *meaning* of the specification of Figure 2 that proofs must be short. A proof in this specification must be a combined record of independent steps, not an interacting record with combined resolutions of ambiguities. To my knowledge, the possibility for this kind of declarative search control in disjunctive modal specifications has not received comment previously. But it seems to me to be one of the most exciting and unique uses for modal logic in representation

$$\frac{\Gamma \to G^{\mu\alpha}, \Delta}{\Gamma \to \Box G^{\mu}, \Delta}\ (\to \Box)$$

$$\frac{\Gamma, G^{\mu\nu} \to \Delta}{\Gamma, \Box G^{\mu} \to \Delta}\ (\Box \to)$$

$$\frac{\Gamma, P^{\mu} \to G^{\mu}, \Delta}{\Gamma \to P \supset G^{\mu}, \Delta}\ (\to\supset)$$

$$\frac{\Gamma \to G^{\mu}, \Delta \qquad \Gamma, P^{\mu} \to \Delta}{\Gamma, G \supset P^{\mu} \to \Delta}\ (\supset\to)$$

$$\frac{\Gamma, A^{\mu} \to \Delta \qquad \Gamma, B^{\mu} \to \Delta}{\Gamma, A \vee B^{\mu} \to \Delta}\ (\vee \to)$$

$$\frac{\Gamma \to A^{\mu}, \Delta \qquad \Gamma \to B^{\mu}, \Delta}{\Gamma \to A \wedge B^{\mu}, \Delta}\ (\to \wedge)$$

$$\Gamma, A^{\mu} \to A^{\mu}, \Delta \quad (Axiom)$$

Figure 3: Six inference figures and the axiom for explicitly-scoped S4. See [Fitting, 1983, Wallen, 1990, Stone, 1999]. The $(\to \Box)$ inference is subject to an eigenvariable condition that α is new. In the $(\Box \to)$ inference, $\mu\nu$ refers to any sequence of terms that extends μ by a suffix ν.

and problem-solving.

1.1.2 Explicit scope and goal-directed search

The second formulation of modal reasoning is illustrated by the sequent figure below:

$$\frac{\Gamma \longrightarrow G^{\mu\alpha}, \Delta}{\Gamma \longrightarrow \Box G^{\mu}, \Delta}\ (\to \Box)$$

Such inferences institute an explicitly-scoped sequent calculus; each formula is tagged with a label indicating the modal context which it describes. These labels are sequences of terms, each of which corresponds to an inference that changes scope. Superscripts are my notation for labels; above, μ labels the scope of the succedent formula $\Box G$. To reason about a generic modal formula, we again introduce a new, nested scope in which just generic information is available; we now label the formula with its new scope. Thus above G is labeled $\mu\alpha$; and α is subject to an eigenvariable condition—it cannot occur elsewhere in the sequent—and so represents a generic possibility. At axioms, the scopes of premises and conclusions must match; therefore modal inferences can dispense with destructive transformation of sequents.

Figure 3 illustrates the other explicitly-scoped S4 sequent inferences that I will draw on in this motivating discussion. Explicitly-scoped proof systems have a long history as *prefixed tableaus*; see [Fitting, 1983, Wallen, 1990] and references therein. Each label sequence can be viewed as representing a possible world in possible-worlds semantics, so for example the $(\to \Box)$ inference figure represents a transition from the world named by μ to a new world $\mu\alpha$ that represents a generic possibility accessible from μ. The more general study of such systems has put them in a new proof-theoretic perspective recently. They are closely related to semantics-based translation systems [Ohlbach, 1991, Nonnengart, 1993] and *labelled deductive systems* [Gabbay, 1996, Basin et al., 1998]. I use the term *explicitly-scoped* from [Stone, 1999] because I continue to emphasize the extent to which the two formulations of reasoning represent the same inferences, just in different ways.

The ability to define explicit scope is intimately connected with the ability to carry out goal-directed proof. I adopt the perspective due to [Miller et al., 1991] that goal-directed proof simply amounts to a specific strategy for constructing sequent calculus deductions. The strategy is first to apply inferences that decompose goals to atoms and then to apply inferences that use a specific program statement to match a specific goal. Proofs that respect this strategy are called *uniform*. On this strategy, logical connectives amount to explicit instructions for search, and this is in fact what lets us view a logical formula concretely as a *program* [Miller et al., 1991].

Unlike other, more procedural characterizations of algorithmic proof, such as [Gabbay, 1992], this view largely abstracts away from the exact state of computations during search. The key questions are purely proof-theoretic. In particular, goal-directed proof is possible in a logic if and only if any theorem has a uniform proof. In systems of structural scope, this is not possible, and we must instead restrict our to inference in specific logical fragments, as described for the intuitionistic case in, e.g., [Miller et al., 1991, Harland, 1994, Harland et al., 2000].[3]

By contrast, systems of explicit scope can be lifted by a suitable analogue to the Herbrand-Skolem-Gödel theorem for classical logic so that *any* pair of unrelated inferences can be interchanged [Kleene, 1951, Wallen, 1990, Lincoln and Shankar, 1994, Stone, 1999]. Thus, unlike systems of structural scope, systems of explicit scope permit *general* goal-directed reasoning. If we adopt Miller's characterization of uniform proof for sequent calculi with multiple conclusions [Miller, 1994, Miller, 1996], then any modal theorem has a uniform proof in a lifted, explicitly-scoped inference system. Put another way, explicitly-scoped inference assimilates modal proof to classical proof, and we know that uniformity is not really a restriction on classical proof [Harland, 1997, Nadathur, 1998]. This is why my investigation emphasizes questions of information-flow, such as modularity and locality, rather than questions of goal-directed proof *per se*.

I will refer to the proof of Figure 4 to illustrate some of the properties of information-flow in goal-directed search. The proof establishes the conclusion

$$F$$

from assumptions

$$A \vee B, C \vee D, A \supset F, C \supset F, (B \wedge D) \supset F$$

First we must get clear on the reasoning Figure 4 represents. The assumptions in this proof again specify two ambiguities, $A \vee B$ and $C \vee D$. In modal terms, these are local ambiguities

[3]A further case of structural control of inference that has attracted particular interest is linear logic, where linear disjunction must be understood to specify synchronization between concurrent processes rather than proof by case analysis; see, e.g., [Andreoli, 1992, Hodas and Miller, 1994, Pym and Harland, 1994, Miller, 1996, Kobayashi et al., 1999]. The investigation of fragments of linear logic remains essential, as linear logic has no analogue of an explicitly-scoped proof system, and so—unlike intuitionistic logic and modal logic—must be understood as a refinement of classical logic rather than an extension to it [Girard, 1993].

$$\boxed{3}\quad \dfrac{\dfrac{\underline{B},\ldots \longrightarrow \ldots,\underline{B} \qquad \underline{D},\ldots \longrightarrow \ldots,\underline{D}}{B,D,\ldots \longrightarrow \underline{B\wedge D}}\,(\to\wedge) \qquad \underline{F},\ldots \longrightarrow \ldots,\underline{F}}{B,D,\underline{(B\wedge D)\supset F},\ldots \longrightarrow \ldots,\underline{F}}\,(\supset\to)$$

$$\boxed{2}\quad \dfrac{\dfrac{\underline{C},\ldots \longrightarrow \ldots,\underline{C} \qquad \boxed{3}}{B,\underline{C\vee D},(B\wedge D)\supset F,\ldots \longrightarrow \ldots,\underline{C},F}\,(\vee\to) \qquad \underline{F},\ldots \longrightarrow \ldots,\underline{F}}{B,C\vee D,\underline{C\supset F},(B\wedge D)\supset F \longrightarrow \ldots,\underline{F}}\,(\supset\to)$$

$$\boxed{1}\quad \dfrac{\dfrac{\underline{A},\ldots \longrightarrow \underline{A},F \qquad \boxed{2}}{\underline{A\vee B},C\vee D,C\supset F,(B\wedge D)\supset F \longrightarrow \underline{A},F}\,(\vee\to) \qquad \underline{F},\ldots \longrightarrow \underline{F}}{A\vee B,C\vee D,\underline{A\supset F},C\supset F,(B\wedge D)\supset F \longrightarrow \underline{F}}\,(\supset\to)$$

Figure 4: A goal-directed proof in which multiple cases are considered. Each case is displayed in a separate block.

that introduce local assumptions; but actually, Figure 4 uses only classical connectives, and this classical reasoning suffices for my discussion here. In Figure 4, the two ambiguities interact to require inference for three separate cases: one case where A is true, one case where C is true, and a final case where B and D are true together. (These cases are laid out separately in Figure 4.) In goal-directed inference, we discover these cases by backward chaining from the main goal F through a series of implications: $A \supset F$, $C \supset F$, and $(B \wedge D) \supset F$.

Now we can describe the structure of the proof of Figure 4 more precisely. The inference is segmented out into three chunks, one for each case. The chunks are indexed to indicate how they should be assembled into a single proof-tree; the chunk indexed $\boxed{3}$ should appear as a subtree where the index $\boxed{3}$ is used in chunk $\boxed{2}$, and that chunk should in turn appear as a subtree where the index $\boxed{2}$ is used in chunk $\boxed{1}$. We could imagine writing out that tree in full—on an ample blackboard! However, the chunks are actually natural units of the proof of Figure 4; they are what Loveland calls *blocks* [Loveland, 1991, Nadathur and Loveland, 1995]. In general, a *block* of a derivation is a maximal tree of contiguous inferences in which the right subtree of any $(\vee \rightarrow)$ inference in the block is omitted. (Check this in Figure 4.) Each block presents reasoning that describes a single case from the specification.

Within blocks, we can trace the progress of goal-directed reasoning, as follows. At each step, our attention is directed to a distinguished goal formula—the current goal—and at most one distinguished program formula—the selected statement. For illustration, these distinguished formulas are underlined in each sequent in Figure 4. Logical operations apply only to distinguished formulas; we first decompose the goal down to atomic formulas, then select a program formula and reason from it to establish the current goal.

There are two things to notice about this derivation. First, we use a *restart* discipline when handling disjunctive case analysis across blocks [Loveland, 1991, Nadathur and Loveland, 1995]. In each new block, the current goal is reset to the original goal F to restart proof search. It is easy to see that it does not suffice, in general, to keep the current goal across blocks; in Figure 4, for example, keeping the current goal would mean continuing to try to prove A after we turn from the case of A to the case of B. The more general restart rule is however complete; in fact, the restart rule is a powerful way to extend a goal-directed proof system to logics where a single proof must sometimes analyze the same goal formula in qualitatively different ways [Gabbay and Reyle, 1984, Gabbay, 1985, Gabbay, 1992].

Second, note when and how newly-assumed disjuncts are used in new blocks. For example, B is assumed in block $\boxed{2}$ but it is not used until block $\boxed{3}$. By contrast, D is assumed in block $\boxed{3}$ and used immediately there. Following [Loveland, 1991], I will refer to any use of a disjunctive premise in the first block of case analysis where it is assumed as a *cancellation*; I will also say that the inference that introduces the disjunct, and the new block it creates, are *canceled*. The proof of Figure 4 cannot be recast in terms of canceled inferences using the sequent rules of Figure 3. Whichever case of $A \vee B$ or $C \vee D$ is treated first cannot be canceled; the second disjunct of the one must wait to be used until the second disjunct of the other is introduced. This is a gap between Loveland's original Near-Horn Prolog interpreter [Loveland, 1991], which requires cancellations, and the generalized reformulation in terms of sequent calculi given in [Nadathur and Loveland, 1995, Nadathur, 1998] and suggested in Figure 4. Loveland suggests that cancellation is just an optimization, but we shall see that modal logic establishes an important proof-theoretic link between cancellation and modularity.

1.1.3 On modular goal-directed proof search

As befits alternative proof methods for the same logic, structurally-scoped systems and explicitly-scoped systems are very close. In fact, in the case of intuitionistic logic, they define not just the same theorems but the same proofs [Stone, 1999]. This correspondence suggests that we use insights about information-flow in structurally-scoped proofs—including the modularity and locality exhibited by Figure 2—to restrict goal-directed proof-search in explicitly-scoped systems.

In fact, we know from [Stone, 1999] that we can sometimes enforce a straightforward requirement of locality in explicitly-scoped inferences, as follows. Assume that we have an explicitly-scoped proof system for a logic with modularity and locality, with an eigenvariable condition on $(\rightarrow \Box)$, and we work in a fragment of logic without negation (this again includes the S4 translation of intuitionistic formulas). Then when we consider a sequent of this form in proof-search:

$$\Gamma \longrightarrow \Delta$$

we apply inferences to a formula P in Γ only when P is labeled with a prefix of a label of a formula in Δ. That is, we can consider inferences on $P^{\mu} \in \Gamma$ only when there is some $G^{\mu\nu} \in \Delta$. The prefix relationship is required for P to eventually contribute to the proof of any Δ formula. For example here:

$$A, B^{\beta}, C^{\gamma}, D^{\delta} \longrightarrow E^{\beta}, F^{\gamma}$$

We can consider A, B^{β} or C^{γ}, but not D^{δ}.

This invariant is weaker than one might want or expect for certain kinds of goal-directed search. Specifically, we have seen that when we use goal-directed search as a model for logic programming, we understand the interpreter to be working on at most one goal and one program statement at a time. In this setting, we would like to require that the program statement must be able to contribute to the current goal. However, we must understand the result of [Stone, 1999] to require only that the current program statement must contribute to *some* goal, not necessarily the *current* one. In addition, we must take into account other, inactive goals, such as the goals that we may potentially restart later in proof search. In the preceding example, even if E^{β} were the current goal, we might have to consider reasoning with C^{γ} because of the possibility of a restart with F^{γ}.

For most inferences, we can rule out their contribution to inactive goals, on independent grounds. Most inferences from assumptions in goal-directed proofs must contribute to the current goal or not at all. But there is one difficult case, which happens also the most meaningful one. This is the case of disjunction itself, where only one disjunct contributes to the current goal. The other disjunct may contribute to some other goal; we will set up a new proof problem by assuming this disjunct and making some inactive goal active. Modularity and locality suggest that we should be able now to select a goal that our newly-assumed disjunct could contribute to. In other words, if the new disjunct is P^{μ} the next goal should take the form $G^{\mu\nu}$ with μ a prefix of $\mu\nu$. Call this a modular restart. The alternative is that there is no relationship of scope between the new disjunct and the next goal.

Modular restarts would be quite powerful. For example, they would allow us to capture the declarative search control illustrated in Figure 2. In an explicitly-scoped goal-directed proof corresponding to Figure 2, the case analysis for $\Box(A \vee B)$ will look like this:

$$\dfrac{\dfrac{\underline{A}^{\alpha},\ldots \longrightarrow \underline{A}^{\alpha},\ldots \qquad B^{\alpha},\ldots \longrightarrow A^{\alpha},\ldots}{\underline{A \vee B}^{\alpha},\ldots \longrightarrow \underline{A}^{\alpha},\ldots}\ \vee\rightarrow}{\underline{\Box(A \vee B)},\ldots \longrightarrow \underline{A}^{\alpha},\ldots}\ \Box\rightarrow$$

With modular restarts, we know we must continue to take A^{α} as the current goal in the right top (B^{α}) subproof. In effect, we know to build a short proof in which ambiguities are considered independently. We can cut down the space for proof search accordingly—for example, there will be no question of introducing the other ambiguity from $\Box(C \vee D)$ in the new modular block. On the other hand, without modular restarts, we are free to reconsider the initial goal $(\Box A) \wedge (\Box C)$ at this stage; in subsequent search we will reconsider both $\Box(A \vee B)$ and $\Box(C \vee D)$. Thus, even though the logic guarantees that ambiguities do not interact in a proof, we still wind up considering interacting ambiguities in proof search.

The main result of this paper is to provide an explicitly-scoped goal-directed proof system in which modular restarts are complete. The proof system has modular restarts because, in the new proof system, any proof can be presented in such a way that all disjunctions are canceled. Each new disjunct P^{μ} therefore contributes to the proof of the restart goal in the current block, and so we know to choose a restart goal $G^{\mu\nu}$ that the new disjunct could contribute to.

It turns out that modular restarts are not automatic; you need to design the policy for disjunctive inference to respect it. Figure 4 already makes the problem clear. How can we enforce cancellations here? The sequent rules seem not to allow it. The new idea is simple actually—to allow a new inference figure for disjunction that considers disjuncts out of their textual order:

$$\dfrac{\Gamma, D^{\mu} \longrightarrow \Delta \qquad \Gamma, C^{\mu} \longrightarrow \Delta}{\Gamma, C \vee D^{\mu} \longrightarrow \Delta}\ \vee\rightarrow *$$

This is the direct analogue of the Near-Horn Prolog inference scheme, which can proceed by

$$\boxed{2'} \quad \dfrac{\underline{C}, \ldots \longrightarrow \ldots, \underline{C} \qquad \underline{F}, \ldots \longrightarrow \ldots, \underline{F}}{B, C, \underline{C \supset F}, \ldots \longrightarrow \ldots, \underline{F}}\,(\supset\rightarrow)$$

$$\boxed{3'} \quad \dfrac{\dfrac{\underline{B}, \ldots \longrightarrow \ldots, \underline{B} \qquad \dfrac{\underline{D}, \ldots \longrightarrow \ldots, \underline{D} \qquad \boxed{2'}}{B, \underline{C \vee D}, C \supset F, \ldots \longrightarrow \ldots, \underline{D}}\,(\vee \rightarrow *)}{B, C \vee D, C \supset F, \ldots \longrightarrow \underline{B \wedge D}}\,(\rightarrow \wedge) \qquad \underline{F}, \ldots \longrightarrow \ldots, \underline{F}}{B, C \vee D, C \supset F, \underline{(B \wedge D) \supset F}, \ldots \longrightarrow \ldots, \underline{F}}\,(\supset\rightarrow)$$

$$\boxed{1} \quad \dfrac{\dfrac{\underline{A}, \ldots \longrightarrow \underline{A}, F \qquad \boxed{3'}}{\underline{A \vee B}, C \vee D, C \supset F, (B \wedge D) \supset F \longrightarrow \underline{A}, F}\,(\vee \rightarrow) \qquad \underline{F}, \ldots \longrightarrow \underline{F}}{A \vee B, C \vee D, \underline{A \supset F}, C \supset F, (B \wedge D) \supset F \longrightarrow \underline{F}}\,(\supset\rightarrow)$$

Figure 5: A reanalysis of the proof of Figure 4 to enforce cancellations. We rewrite block $\boxed{3'}$, in which B is canceled, to use the new disjunctive inference figure; block $\boxed{3'}$ thus becomes the second block after $\boxed{1}$. At the same time, we introduce a simplified block $\boxed{2'}$ which uses the assumption C, without disjunction at all.

matching any of the heads of a disjunctive clause at any time [Loveland, 1991]. The new sequent rule will allow us to reanalyze the constitution of higher blocks so that, wherever we use the new disjunct in the original proof, we can always reanalyze it as part of the current block. Figure 5 demonstrates this reanalysis for Figure 4. In fact, demonstrating the generality of such reanalysis will prove to be quite involved. Explicitly-scoped inferences with an eigenvariable condition give blocks in modal proofs an inherently hierarchical structure, because of the different modal scopes that are introduced and the local assumptions that are made. Loveland's construction for cancellations, by contrast, assumes that the structure of blocks is flat. Instead, we must use the natural tools of the sequent calculus to develop suitable constructions for reanalyzed inferences.

1.2 The results and their context

The problem sketched in Section 1.1 is a pure problem of modal proof. Accordingly, all the proof systems I consider will describe sound and complete inference under the usual Kripke semantics for modal logic [Kripke, 1963, Fitting, 1983]. I will not consider interactions of disjunction with negation-by-failure and other operational features of of logic programming proof-search systems. For such issues in disjunctive logic programming, see for example [Lobo et al., 1992]. Nor will I attempt to describe a minimal model or fixed-point construction in which exactly the consequences of a modal program hold, as in [Orgun and Wadge, 1992].

Moreover, my interest is in specific fragments of specific modal logics in particular. *Modularity* and *locality* allow consideration of the logics T, K and K4 in addition to S4, but are not compatible with such logics as S5, temporal logics with symmetric past and future operators [Gabbay, 1987], the logic of context of [McCarthy and Buvač, 1994] or the modal logic of named addresses of

[Kobayashi et al., 1999]. For example, in S5, if $\Box A$ is true at any world, then $\Box A$ is true at all worlds; thus the logic prohibits making such an assumption locally. To see the problem, observe, for example, that $\Box(\Box A \supset B) \vee \Box A$ is a theorem of S5. Modal proofs in such cases require global restarts [Gabbay and Olivetti, 1998]. *Locality* further rules out logical fragments with possibility or negation. Such fragments can be used to pose goals about that access otherwise local assumptions, as in the theorem $\Box(A \supset B) \vee \Diamond A$ of all normal modal logics. (Goal-directed proof of this theorem also involves a global restart.) Moreover, such fragments make it more difficult to enforce modularity as well, since they do not permit an eigenvariable condition at $(\rightarrow \Box)$ inferences in goal-directed proofs. My investigation therefore sticks closely to the treatments of logical modularity and locality originally explored in [Miller, 1989, Giordano and Martelli, 1994]. Indeed, I continue to restrict implications and universal quantifiers in goals to strict statements of the form $\Box(P \supset G)$ and $\Box\forall xG$.

The basic strategy that I adopt is to start with a relatively straightforward proof system, and gradually to narrow the formulation of its inference rules—preserving soundness and completeness with respect to the underlying semantics—until we have a proof system, SCLP, with the desired characteristics, namely goal-directed search and modular restarts. I have been particularly influenced by Lincoln and Shankar's presentation of proof-theoretic results in terms of simple transformations among successive proof systems [Lincoln and Shankar, 1994]; and by Andreoli's construction of focusing sequent calculi that embody the discipline of goal-directed proof directly in the form of inference figures [Andreoli, 1992].

However, the correct design of the final proof system requires a variety of proof-theoretic ideas about logic programming to be extended, strengthened, and combined with proof-theoretic results about modal logic in a novel way. To describe logic programming, we start with the idea of uniform proof search described in [Miller et al., 1991] and extended to multiple-conclusion calculi in [Miller, 1994]. To derive a uniform proof system in the presence of indefinite information in assumptions, however, we can no longer use the familiar quantifier rules used in previous logic programming research, which simply introduce fresh parameters; we must apply a generalization of Herbrand's Theorem [Lincoln and Shankar, 1994] and work with quantifier rules that introduce structured terms. The calculus of Herbrand terms, SCL, lifts the explicitly-scoped proof systems considered in Section 1.1.2 and [Fitting, 1983, Wallen, 1990]. The key property of SCL is that inferences can be freely interchanged. This allows arbitrary proofs to be transformed easily into uniform proofs.

The *modular* behavior of this uniform system depends on the further proof-theoretic analyses of path-based sequent calculi adapted, in part, from [Stone, 1999]. These analyses establish that path representations enforce modularity and locality in the uses of formulas in proofs, even with otherwise classical reasoning. Hence, although path-based calculi obscure the natural modularity of modal inference, they do not eliminate it. I finish with a streamlined uniform proof system that takes advantage of these results; as a consequence, proof search in this calculus can dynamically exploit the local use of modular assumptions.

The justification of this new proof system makes much of a strategy originally due to [Kleene, 1951], in which the inferences in a proof are reordered so as to satisfy a global invariant. The strategy achieves termination despite generous copying and deepening of inferences by a judicious choice of transformations within a double induction. In our cases, these transformations are guided by the constraints of uniform proof, and by the cancellations of disjunctive assumptions that we know we must maintain in proofs, to achieve modularity. This provides an analogue of Loveland's transformations on restart proofs [Loveland, 1991] in the sequent calculus setting.

Of course, modal logic is not just a modular logic. Modal logic provides a general, declarative formalism for specifying change over time, the knowledge of agents, and other special-purpose domains [Prior, 1967, Hintikka, 1971, Schild, 1991]. Goal-directed systems for modal proof are often motivated by such specifications [Fariñas del Cerro, 1986, Debart et al., 1992, Baldoni et al., 1993, Baldoni et al., 1996]. In generalizing goal-directed modal proof to indefinite specifications, SCLP can play an important role in applying modal formalisms to planning, information-gathering and communication [Stone, 1998a, Stone, 2000]. Even when content, not modularity, is primary, the modular treatment of disjunction limits the size of proofs and the kinds of interactions that must be considered in proof search. Such constraints are crucial to the use of logical techniques in applications that require automatic assessment of incomplete information, such as planning and natural language generation. The interest of these more general applications helps explain why I pursue this investigation in the full first-order language.

1.3 Outline

The structure of the rest of this paper is as follows. I begin by presenting first-order multi-modal logic in Section 2. I consider syntax (Section 2.1), semantics (Section 2.2), and finally proof (Section 2.3); I describe the explicitly-scoped Herbrand proof system for modal logic that is my starting point. Section 2.4 shows that this calculus offers a suitable framework for goal-directed proof because uniform proof search in this calculus is complete.

Section 3 describes and justifies a modular goal-directed proof system, as advertised in Section 1.1. I introduce the calculus itself in Section 3.1, along with key definitions and examples. Then in Sections 3.2–3.4 I outline how this sequent calculus is derived in stages from the calculus of Section 2. Full details are provided in an appendix.

Finally, Section 4 offers a broader assessment of these results. I consider some further optimizations that the new sequent calculus invites in Section 4.1, and briefly conclude in Section 4.2 with some applications of first-order multi-modal inference that the new sequent calculus suggests.

2 First-order multi-modal deduction

I begin by supporting the informal presentation of first-order multi-modal logic from Section 1 more explicitly. I will adopt a number of techniques that are individually quite familiar. I allow an arbitrary number of modal operators and a flexible regime for relating different modal operators to one another, following many applied investigations [Debart et al., 1992, Baldoni et al., 1993, Baldoni et al., 1996, Baldoni et al., 1998b]. I use prefix terms for worlds and sequent calculus inference, following the comprehensive treatment of the first-order modal logic using prefix terms and analytic tableaux (or, seen upside-down, in the cut-free sequent calculus) of [Fitting and Mendelsohn, 1998]. I factor out reasoning about accessibility into side conditions on inference rules, similar to the proof-theoretic view of [Basin et al., 1998], in which reasoning about accessibility and boolean reasoning are clearly distinguished. And I use Herbrand terms to reason correctly about parameterized instances of formulas, avoiding the usual eigenvariable condition on quantifier (and modal) rules, as in [Lincoln and Shankar, 1994].

Though the techniques are routine, the combination is still somewhat unusual. Research in modal logic—whether the investigation is more mathematical [Goré, 1992, Massacci, 1998b, Massacci, 1998a, Goré, 1999] or primarily concerns algorithms for proof search [Otten and Kreitz, 1996, Beckert and Goré, 1997, Schmidt, 1998]—is dominated by the study of the propositional logic of a single modal operator (or accessibility relation). Moreover, researchers who have investigated modal logic in a first-order setting have tended to jump directly into a discussion of particular theorem-proving strategies, particularly resolution [Jackson and Reichgelt, 1987, Wallen, 1990, Catach, 1991, Frisch and Scherl, 1991, Auffray and Enjalbert, 1992, Nonnengart, 1993, Ohlbach, 1993].

2.1 Syntax

Our language depends on a *signature* including a suitable set of atomic constants C (and suitable predicate symbols and modalities). We then consider program statements of the syntactic category $D(C)$ and goals of the category $G(C)$ defined recursively as in (1); we refer to the union of these two languages as $L(C)$. (1) makes explicit the conditions observed in Section 1.2; there is no possibility or negation, and universal and hypothetical goals must be modal.

(1) $$G ::= A \mid [\textsc{m}]G \mid G \wedge G \mid G \vee G \mid [\textsc{m}](\forall x G) \mid \exists x G \mid [\textsc{m}](D \supset G)$$
$$D ::= A \mid [\textsc{m}]D \mid D \wedge D \mid D \vee D \mid \forall x D \mid \exists x D \mid G \supset D$$

In (1), A schematizes an atomic formula; atomic formulas take the form $p_i(a_1, \ldots, a_k)$ where p_i is a predicate symbol of arity k and each a_i is either a variable or an atomic constant in the set C. We assume some initial non-empty set of constants $CONST$. But it will be convenient to consider languages in which a countably infinite number of *parameters* are included in the language to supplement the symbols in CONST.

In (1), $[\textsc{m}]$ schematizes a modal operator of necessity; intuitively, such modal operators allow a specification to manipulate constrained sources of information. That is, a program statement $[\textsc{m}]D$ explicitly indicates that D holds in the constrained source of information associated with the operator $[\textsc{m}]$. Conversely, a goal $[\textsc{m}]G$ can be satisfied only when G is established by using information from the constrained source associated with $[\textsc{m}]$.

We will work in a multi-modal logic, in which any finite number m of distinct necessity operators or *modalities* may be admitted. (Necessity operators will also be written as $\Box$ or $\Box_i$.) In addition to ordinary program statements, a specification may contain any of the following axiom schemes describing the modalities to be used in program statements and goals:

(2)
$\Box_i p \supset p$ *veridicality* (VER)
$\Box_i p \supset \Box_i \Box_i p$ *positive introspection* (PI)
$\Box_i p \supset \Box_j p$ *inclusion* (INC)

These axioms describe the nature of the information that an operator provides, and spell out relationships among the different sources of information in a specification. (VER) is needed for information that correctly reflects the world; (PI), for information that provides a complete picture of how things might be; and (INC), for a source of information, j, that elaborates on information from another source, i. Because we use this explicit axiomatization, we can take the names of the modal operators as arbitrary.

We appeal to the usual notions of *free* and *bound* occurrences of variables in formulas; we likewise invoke the *depth* of a formula—the largest number of nested logical connectives in it.

2.2 Semantics

As is standard, we describe the models for the modal language in two steps. The first step is to set up *frames* that describe the structure of any model; a full model can then be obtained by combining a frame with a way of assigning interpretations to formulas in a frame.

Definition 1 (Frame) *A* frame *consists of a tuple* $\langle \mathbf{w}, \mathcal{R}, \mathcal{D} \rangle$ *where:* $\mathbf{w}$ *is a non-empty set of* possible worlds*;* $\mathcal{R}$ *names a family of m binary* accessibility *relations on* $\mathbf{w}$*, a relation* $\mathcal{R}_i$ *for each modality i; and* $\mathcal{D}$ *is a* domain function *mapping members of* $\mathbf{w}$ *to non-empty sets.*

Within the frame $\mathcal{F}$, the function $\mathcal{D}$ induces a set $\mathcal{D}(\mathcal{F})$, called the *domain of the frame*, as $\cup\{\mathcal{D}(w) \mid w \in \mathbf{w}\}$. In order to simplify the treatment of constant symbols, it is also convenient to define a set of objects that all the domains of the different possible worlds have in common, the *common domain of the frame* $\mathcal{F}$: $\mathcal{C}(\mathcal{F}) = \cap\{\mathcal{D}(w) \mid w \in \mathbf{w}\}$. We effectively insist that $\mathcal{C}(\mathcal{F})$ be non-empty as well, since CONST is non-empty and each symbol in CONST must be interpreted by an element of $\mathcal{C}(\mathcal{F})$.

The intermediate level of frames is useful in characterizing the meanings of modal operators and modal quantification. In particular, simply by putting constraints on $\mathcal{R}_i$ or on $\mathcal{D}$ at the level of frames, we can obtain representative classes of models in which certain general patterns of inference are validated. The constraints we will avail ourselves of are introduced in Definition 2.

Definition 2 *Let* $\langle \mathbf{w}, \mathcal{R}, \mathcal{D} \rangle$ *be a* frame. *We say the frame is:*

- reflexive *at i if* $w\mathcal{R}_i w'$ *for every* $w \in \mathbf{w}$*;*
- transitive *at i if, for any* $w, w'' \in \mathbf{w}$*,* $w\mathcal{R}_i w''$ *whenever there is a* $w' \in \mathbf{w}$ *such that* $w\mathcal{R}_i w'$ *and* $w'\mathcal{R}_i w''$*;*
- narrowing *from i to j if* $w\mathcal{R}_j w'$ *implies* $w\mathcal{R}_i w'$ *for all* $w, w' \in \mathbf{w}$*;*
- increasing domain *if for all* $w, w' \in \mathbf{w}$*,* $\mathcal{D}(w) \subseteq \mathcal{D}(w')$ *whenever there is some accessibility relationship* $w\mathcal{R}_i w'$*.*

Our scheme for using the constraints of Definition 2 depends on establishing a regime for the *m* modalities in the language, describing the inferences that should relate them. The regime is defined as follows.

Definition 3 (Regime) *A* regime *is a tuple* $\langle \mathcal{A}, \mathcal{N}, \mathcal{Q} \rangle$*, where:* $\mathcal{A}$ *is a function mapping each modality into one of the symbols K, K4, T and S4;* $\mathcal{N}$ *is a (strict) partial order on the modalities; and* $\mathcal{Q}$ *is the symbol* increasing.

The reader will recognize the symbols in the image of $\mathcal{A}$ as the classic names for modal logics of a single modality. *S*4 is for modalities that are subject to (PI) and (VER). *T* is for modalities that are subject just to (VER). *K*4 is for modalities that are subject just to (PI). *K* is modalities subject to neither axiom. The interactions specified by (INC) are determined by the partial order on modalities: $j \leq i$ when $\Box_i p \supset \Box_j p$. This meaning for these symbols can be enforced by considering only frames that *respect* the given *regime*.

Definition 4 (Respect) *Let $\mathcal{F} = \langle \mathbf{w}, \mathcal{R}, \mathcal{D} \rangle$ be a frame, and let $\mathcal{S} = \langle \mathcal{A}, \mathcal{N}, \mathcal{Q} \rangle$ be a regime. We say $\mathcal{F}$* respects $\mathcal{S}$ *whenever the following conditions are met for all modalities i and j:*

- *If $\mathcal{A}(i)$ is T or S4 then $\mathcal{R}_i$ is reflexive.*
- *If $\mathcal{A}(i)$ is K4 or S4 then $\mathcal{R}_i$ is transitive.*
- *If $j \leq i$ according to $\mathcal{N}$ then $\mathcal{F}$ is narrowing from i to j.*
- *If $\mathcal{Q}$ is* increasing, *then $\mathcal{F}$ is increasing domain.*

From now on, we assume that some regime $\mathcal{S} = \langle \mathcal{A}, \mathcal{N}, \mathcal{Q} \rangle$ is fixed, and restrict our attention to frames that respect $\mathcal{S}$. Informally, now, a model consists of a frame together with an interpretation.

Definition 5 (Interpretation) *$\mathcal{J}$ is an* interpretation *in a frame $\langle \mathbf{w}, \mathcal{R}, \mathcal{D} \rangle$ if $\mathcal{J}$ satisfies these two conditions:*

1. *$\mathcal{J}$ assigns to each n-place relation symbol p_i and each possible world $w \in \mathbf{w}$ some n-ary relation on the domain of the frame $\mathcal{D}(\mathcal{F})$.*
2. *$\mathcal{J}$ assigns to each constant symbol c some element of the common domain of the frame $\mathcal{C}(\mathcal{F})$.*

Thus we can define a *model* over $\mathcal{S}$ thus:

Definition 6 (Model) *A first-order k*-modal model *over a regime $\mathcal{S}$ is a tuple $\langle \mathbf{w}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle$ where $\langle \mathbf{w}, \mathcal{R}, \mathcal{D} \rangle$ is a frame that respects $\mathcal{S}$ and $\mathcal{J}$ is an interpretation in $\langle \mathbf{w}, \mathcal{R}, \mathcal{D} \rangle$.*

To define truth in a model, we need the usual notion of assignments and variants:

Definition 7 (Assignment) *Let $\mathcal{M} = \langle \mathbf{w}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle$ be a model (that respects the regime $\mathcal{S}$). An* assignment *in $\mathcal{M}$ is a mapping g that assigns to each variable x some member $g(x)$ of the domain of the frame of the model $\mathcal{D}(\langle \mathbf{w}, \mathcal{R}, \mathcal{D} \rangle)$.*

In proofs, we interpret formulas not just in the ordinary language $L(C)$ with a given set of modalities, relations, constants and variables, but in an expanded language $L(C \cup P)$ which also includes a set P of first-order *parameters*; we will want to use the same models for this interpretation. Over $L(C \cup P)$, we suppose that an assignment in $\mathcal{M}$ also assigns some member $g(p)$ of the domain of the frame of $\mathcal{M}$ to each parameter p in P.

Definition 8 (Variants) *Let g and g' be two assignments in a model $\mathcal{M} = \langle \mathbf{w}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle$; g' is an x*-variant of g at a world $w \in \mathbf{w}$ *if g and g' agree on all variables except possibly for x and $g'(x) \in \mathcal{D}(w)$.*

Definition 9 (Truth in a model) *Let $\mathcal{M} = \langle \mathbf{w}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle$ be a model. Then the formula A is* true at world w of model $\mathcal{M}$ on assignment *g—written $\mathcal{M}, w \Vdash_g A$—just in case the clause below selected by syntactic structure of A is satisfied:*

- *A is $p_i(t_1, \ldots, t_n)$: Then $\mathcal{M}, w \Vdash_g A$ just in case $\langle e_1, \ldots e_n \rangle \in \mathcal{J}(p_i, w)$, where for each t_i, e_i is $\mathcal{J}(t_i)$ if t_i is a constant and $g(t_i)$ otherwise.*

- *A is $B_1 \wedge B_2$: Then $\mathcal{M}, w \Vdash_g A$ just in case both $\mathcal{M}, w \Vdash_g B_1$ and $\mathcal{M}, w \Vdash_g B_2$.*
- *A is $B_1 \vee B_2$: Then $\mathcal{M}, w \Vdash_g A$ just in case either $\mathcal{M}, w \Vdash_g B_1$ or $\mathcal{M}, w \Vdash_g B_1$.*
- *A is $\Box_i B$: Then $\mathcal{M}, w \Vdash_g A$ just in case for every $w' \in \mathbf{w}$, if $w\mathcal{R}_i w'$ then $\mathcal{M}, w' \Vdash_g B$.*
- *A is $\forall x B$: Then $\mathcal{M}, w \Vdash_g A$ just in case for every x-variant g' of g at w, $\mathcal{M}, w \Vdash_{g'} B$.*
- *A is $\exists x B$: Then $\mathcal{M}, w \Vdash_g A$ just in case there is some x-variant g' of g at w with $\mathcal{M}, w \Vdash_{g'} B$.*

By a *sentence* we mean a formula of $L(\mathrm{CONST})$ in which no variables occur free. For any sentence A, model $\mathcal{M}$ and world w of $\mathcal{M}$, a simple induction on depth guarantees that $\mathcal{M}, w \Vdash_g A$ for some assignment g in $\mathcal{M}$ exactly when $\mathcal{M}, w \Vdash_g A$ for all assignments g in $\mathcal{M}$. In this case, we can write simply $\mathcal{M}, w \Vdash A$ and say that A is *true in $\mathcal{M}$ at w*.

Definition 10 (Valid) *Let A be a sentence and $\mathcal{M} = \langle \mathbf{w}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle$ be a model. A is* valid in $\mathcal{M}$ *if for every world $w \in \mathbf{w}$, $\mathcal{M}, w \Vdash A$. A is* valid *(on the regime $\langle \mathcal{A}, \mathcal{N}, \mathcal{Q} \rangle$) if A is valid in any model $\mathcal{M}$ that respects the regime.*

2.3 Proof theory

We now present our basic deductive system—a cut-free path-based sequent calculus for multi-modal deduction which uses Herbrand terms to reason correctly about parameterized instances of formulas. Since this calculus represents our basic *lifted sequent calculus* for modal logic, we refer to it as SCL here. Our starting point is Theorem 1 that SCL provides a sound and complete characterization of valid formulas.

SCL has the advantage that inferences can be freely interchanged, allowing arbitrary proofs to be transformed easily into goal-directed proofs; we show in Theorem 2, presented in Section 2.4, how to obtain goal-directed proofs in this calculus. The very same flexibility of inference, however, means that this calculus neither respects nor represents the potential of modal inference to give proofs an explicitly modular structure.

The basic constituent in SCL is a *tracked, prefixed formula*. The formulas extend the basic languages $D(C)$ and $G(C)$ of definitions and goals defined in (1) by allowing additional terms—representing arbitrary witnesses of first order quantifiers, and arbitrary transitions of modal accessibility among possible worlds—to be introduced into formulas for the purposes of proof. We begin by assuming two countable sets of symbols: a set H of *first-order Herbrand functions* and Υ of *modal Herbrand functions*. We can now define sets P_H of *first-order Herbrand terms*, κ_Υ of *modal Herbrand terms*, and $\Pi(\kappa_\Upsilon)$ of *Herbrand prefixes* by mutual recursion:

Definition 11 (Herbrand terms and prefixes) *Assume that t_0 is a Herbrand prefix and let $t_1, \ldots, t_n$ be a sequence (possibly empty), where each t_i is either an element of C, a first-order Herbrand term, or a Herbrand prefix. Then if h is a first-order Herbrand function then $h(t_0, t_1, \ldots, t_n)$ is a* first-order Herbrand term. *If η is a modal Herbrand function then $\eta(t_0, t_1, \ldots, t_n)$ is a* modal Herbrand term. *A* Herbrand prefix *is any finite sequence of modal Herbrand terms.*

The rationale behind the use of a Herbrand term $h(X)$ at an existential inference R goes like this. At existential inferences, we need to reason about a generic individual. We need to have a suitable representation for a generic individual for R. Regardless of the order in which inferences are applied in a sequent deduction, there will be some parameters that must occur in the sequent where R applies. For example, some parameters must appear here as a result of the instantiations that must take place in deriving the formula to which R applies. We must be sure that the individual we introduce for R is different from all these parameters. The terms X which are supplied as an argument to the Herbrand term $h(X)$ identify these parameters indirectly. The structure $h(X)$ therefore serves as a placeholder for a new parameter that could be chosen to be different from each of the terms in X. The structure $h(X)$ thus packs all the information required to allow the inferences in the proof to be reordered and an appropriate parameter chosen so that the inference at R is truly generic.

In modal deduction, of course, we need generic individuals at modal inferences as well as existential ones. Modal Herbrand inference therefore requires that we introduce Herbrand terms to describe transitions among possible worlds and Herbrand prefixes to name possible worlds, in addition to introducing first-order Herbrand terms to represent first-order parameters. In this case, the arguments X to Herbrand terms must mix first-order Herbrand terms and Herbrand prefixes, since logical formulas can encode dependencies among first-order and modal parameters.

A *prefixed formula* is now an expression of the form A^μ with A a formula and μ a Herbrand prefix—we use $D(C \cup P_H)^{\Pi(\kappa_\Upsilon)}$ and $G(C \cup P_H)^{\Pi(\kappa_\Upsilon)}$ to refer to prefixed definitions and prefixed goals. For Herbrand calculi, formulas must also be *tracked* to indicate the sequence of instantiations that has taken place in the derivation of the formula.

Definition 12 (Tracked expressions) *If E denotes the expressions of some class, then the* tracked expressions *of that class are expressions of the form e_I where e is an expression of E and I is a finite sequence (possibly empty) of elements of $C \cup P_H \cup \Pi(\kappa_\Upsilon)$.*

We say that a tracked expression e_I *tracks* a term t just in case t occurs as a subterm of some term in I.

In order to reason correctly about multiple modal operators, we need to keep track of the kinds of accessibility that any modal transition represents. To endow the system with correct first-order reasoning on increasing domains, we also need to keep track of the worlds where first-order terms are introduced. We use the following notation to record these judgments: $\mu/\nu : i$ indicates that world ν is accessible from world μ by the accessibility relation for modality i; and $t : \mu$ indicates that the entity associated with term t exists at world μ.

It is convenient to keep track of this information by anticipating the restricted reasoning required for our fragment $L(C)$ and exploiting the structure of Herbrand terms, as follows. It is clear that there are countably many first-order Herbrand terms, Herbrand prefixes, and formulas in $L(C \cup P_H)$. We can therefore describe a correspondence as follows. If A is a formula of the form $\forall x B$ or $\exists x B$ and u is a natural number, we define a corresponding first-order Herbrand function h_A^u so that each first-order Herbrand function is h_A for some A and no first-order Herbrand function is h_A^u and h_B^v for distinct A and B or distinct u and v. Likewise, if A is a formula of the form $\Box_i B$ and u is a natural number, we define a corresponding modal Herbrand function η_A^u so that each modal Herbrand function is η_A^u for some A and no modal Herbrand function is η_A^u and η_B^v for distinct A and B or distinct u and v. (Indexing Herbrand functions by natural numbers means that adapting a Herbrand proof to respect an eigenvariable condition can be as simple as reindexing its Herbrand functions.) Now we have:

Definition 13 (Herbrand typings) *A* Herbrand typing for the language $L(C \cup P_H)$ *(under a correspondence as just described) is a set* Ξ *of statements, each of which takes one of two forms:*

1. $\mu/\mu\eta : i$ *where:* μ *is a Herbrand prefix and* η *is a modal Herbrand term of the form* $\eta^u_A(\mu,\ldots)$ *and* A *is* $\Box_i B$.

2. $t : \mu$ *where* t *is a first-order Herbrand term of the form* $h(\mu,\ldots)$.

A sequence of modal and first-order Herbrand terms X *determines a Herbrand typing* Ξ_X*, consisting of the appropriate* $\mu/\mu\eta : i$ *for each modal Herbrand term* η *that occurs in* X *(possibly as a subterm) and the appropriate* $h : \mu$ *for each first-order Herbrand term* h *that occurs in* X *(possibly as a subterm).*

Definition 14 (Typings) *Suppose that* Ξ *is a Herbrand typing over a language* $L(C \cup P)^{\Pi(\kappa)}$*, and that* $S = \langle A, N, increasing\rangle$ *is a modal regime. We define the relation that* E *is a* derived typing *from* Ξ *with respect to* S*, written* $S,\Xi \triangleright E$*, as the smallest relation satisfying the following conditions:*

- (K). $S,\Xi \triangleright \mu/\nu : i$ *if* $\mu/\nu : i \in \Xi$.
- (T). $S,\Xi \triangleright \mu/\mu : i$ *if* $A(i)$ *is T or S4, and* μ *occurs in* Ξ.
- (4). $S,\Xi \triangleright \mu/\nu : i$ *if* $\mu/\mu' : i \in \Xi$, $S,\Xi \triangleright \mu'/\nu : i$*, and* $A(i)$ *is K4 or S4.*
- (Inc). $S,\Xi \triangleright \mu/\nu : i$ *if* $S,\Xi \triangleright \mu/\nu : j$ *and* $j \leq i$ *according to* N.
- (V). $S,\Xi \triangleright t : \mu$ *if* $t : \mu \in \Xi$.
- (I). $S,\Xi \triangleright t : \nu$ *if* $S,\Xi \triangleright \mu/\nu : i$ *for some* i *and* $S,\Xi \triangleright t : \mu$.

Inspection of these rules shows that $S,\Xi \triangleright \mu/\nu : i$ only if ν and μ occur in Ξ. Moreover, given these rules, an easy induction on the length of typing derivations gives that $S,\Xi \triangleright \mu/\nu : i$ only if $\nu = \mu\nu'$ for some prefix ν'. Thus, suppose that $S,\Xi \triangleright \mu/\nu : i$ for some Herbrand typing Ξ: each step in the derivation must concern some prefix of ν and thus $S,\Xi_\nu \triangleright \mu/\nu : i$. These invariants permit some simplifications in reasoning in the fragment $L(C \cup P)$ over more expressive modal regimes containing other modal operators and other uses of connectives.

These definitions allow us to describe the modal Herbrand sequent calculus precisely. This calculus, SCL, is given in Definition 15. Note that for this fragment of modal logic, it suffices to consider sequents of the form $\Delta \longrightarrow \Gamma$, where Δ is a multiset of prefixed definitions (from $D(C \cup P_H)^{\Pi(\kappa_\Upsilon)}$), and Γ is a multiset of prefixed goals (from $G(C \cup P_H)^{\Pi(\kappa_\Upsilon)}$). Note also that $S,\Xi \triangleright \mu/\nu : i$ only if ν is of the form $\mu\nu'$.

Definition 15 (Herbrand sequent calculus) *For basic first-order multi-modal Herbrand deductions in our fragment over a regime* S*, we will use the sequent rules defined here, which comprise the system SCL. The rules consist of an axiom rule and recursive rules—each recursive rule relates a* base *sequent below to one or more* spur *sequents above; it applies to the base in virtue of an occurrence of a distinguished tracked, prefixed formula in the sequent; we refer to this as the* principal expression *or simply the* principal *of the inference. Similarly, each of the sequent rules* *introduces new expressions onto each spur, which we refer to as the* side expressions *of the rule. We will also refer to the two named expression occurrences at axioms as the* principal expressions *or* principals *of the axiom. Now we have:*

1. *axiom—A atomic:*

$$\Delta, A^{\mu}_{X} \longrightarrow \Gamma, A^{\mu}_{Y}$$

2. *conjunctive:*

$$\frac{\Delta, A \wedge B^{\mu}_{X}, A^{\mu}_{X}, B^{\mu}_{X} \longrightarrow \Gamma}{\Delta, A \wedge B^{\mu}_{X} \longrightarrow \Gamma}\ (\wedge \rightarrow)$$

$$\frac{\Delta \longrightarrow \Gamma, A \vee B^{\mu}_{X}, A^{\mu}_{X}, B^{\mu}_{X}}{\Delta \longrightarrow \Gamma, A \vee B^{\mu}_{X}}\ (\rightarrow \vee)$$

$$\frac{\Delta, A^{\mu}_{X} \longrightarrow \Gamma, A \supset B^{\mu}_{X}, B^{\mu}_{X}}{\Delta \longrightarrow \Gamma, A \supset B^{\mu}_{X}}\ (\rightarrow \supset)$$

3. *disjunctive:*

$$\frac{\Delta \longrightarrow \Gamma, A \wedge B^{\mu}_{X}, A^{\mu}_{X} \qquad \Delta \longrightarrow \Gamma, A \wedge B^{\mu}_{X}, B^{\mu}_{X}}{\Delta \longrightarrow \Gamma, A \wedge B^{\mu}_{X}}\ (\rightarrow \wedge)$$

$$\frac{\Delta, A \vee B^{\mu}_{X}, A^{\mu}_{X} \longrightarrow \Gamma \qquad \Delta, A \vee B^{\mu}_{X}, B^{\mu}_{X} \longrightarrow \Gamma}{\Delta, A \vee B^{\mu}_{X} \longrightarrow \Gamma}\ (\vee \rightarrow)$$

$$\frac{\Delta, A \supset B^{\mu}_{X} \longrightarrow A^{\mu}_{X}, \Gamma \qquad \Delta, A \supset B^{\mu}_{X}, B^{\mu}_{X} \longrightarrow \Gamma}{\Delta, A \supset B^{\mu}_{X} \longrightarrow \Gamma}\ (\supset \rightarrow)$$

4. *possibility—where* η *is* $\eta^{u}_{\Box_i A}(\mu, X)$ *for some u:*

$$\frac{\Delta \longrightarrow \Gamma, \Box_i A^{\mu}_{X}, A^{\mu\eta}_{X,\mu\eta}}{\Delta \longrightarrow \Gamma, \Box_i A^{\mu}_{X}}\ (\rightarrow \Box)$$

5. *necessity—subject to the side condition* $S, \Xi_{\nu} \rhd \mu/\mu\nu : i$*:*

$$\frac{\Delta, \Box_i A^{\mu}_{X}, A^{\mu\nu}_{X,\mu\nu} \longrightarrow \Gamma}{\Delta, \Box_i A^{\mu}_{X} \longrightarrow \Gamma}\ (\Box \rightarrow)$$

6. *existential—subject to the side condition that h is* $h^{u}_{B}(\mu, X)$ *for* B^{μ}_{X} *the principal of the rule (either* $\exists x A$ *or* $\forall x A$*) and some u:*

$$\frac{\Delta, \exists x A^{\mu}, A[h/x]^{\mu}_{X,h} \longrightarrow \Gamma}{\Delta, \exists x A^{\mu}_{X} \longrightarrow \Gamma}\ (\exists \rightarrow) \quad \frac{\Delta \longrightarrow \Gamma, \forall x A^{\mu}_{X}, A[h/x]^{\mu}_{X,h}}{\Delta \longrightarrow \Gamma, \forall x A^{\mu}_{X}}\ (\rightarrow \forall)$$

7. *universal—subject to the side condition* $S, \Xi_{t,\mu} \rhd t : \mu$*:*

$$\frac{\Delta, \forall x A^{\mu}_{X}, A[t/x]^{\mu}_{X,t} \longrightarrow \Gamma}{\Delta, \forall x A^{\mu}_{X} \longrightarrow \Gamma}\ (\forall \rightarrow) \quad \frac{\Delta \longrightarrow \Gamma, \exists x A^{\mu}_{X}, A[t/x]^{\mu}_{X,t}}{\Delta \longrightarrow \Gamma, \exists x A^{\mu}_{X}}\ (\rightarrow \exists)$$

A S-proof or S-derivation for a sequent $\Delta \longrightarrow \Gamma$ is a tree built by application of these inference figures (in such a way that any side conditions are met for regime S), with instances of the axiom as leaves and with the sequent $\Delta \longrightarrow \Gamma$ at the root. A tree similarly constructed except for containing some arbitrary sequent S as a leaf is a *derivation from S*.

I state the correctness theorem for this proof theory in a way that highlights the continuity with previous work on modal logic, particularly [Fitting, 1983].

Theorem 1 (Soundness and Completeness) *Suppose there is an S-proof for a sequent $\longrightarrow A$. Then A is valid. Conversely, if there is no S-proof for the sequent $\longrightarrow A$ then there is a model $\mathcal{M}$ (that respects S) and world w such that $\mathcal{M}, w \not\Vdash A$.*

I merely sketch a proof here, which involves simply applying the standard techniques of [Fitting, 1983, Lincoln and Shankar, 1994]. It is convenient to prove an intermediate result, using slightly modified sequent calculus SCE which imposes an eigenvariable condition on the possibility and existential rules—*u* must be new. We can show the soundness of SCE by adapting the arguments presented in [Fitting, 1983, 2.3] and [Fitting and Mendelsohn, 1998, 5.3]. Meanwhile, we can follow [Fitting, 1983] in developing the completeness argument in terms of *analytic consistency properties* for the modal language. This argument can be seen as a formalization of the motivation for sequent calculi in the systematic search for models. Now, modal formulas may be satisfied only in infinite models, so the completeness theorem effectively requires us to consider infinite sequences of applications of sequent rules. In moving to infinite sets in this way, we must formally move from deductions, viewed as syntactic objects, to a more abstract, algebraic characterization of sets of modal formulas.

We can now establish the correctness of SCL by syntactic methods, which relate SCL proofs to SCE proofs. Suppose Γ and Δ contains sentences of $L(\text{CONST})$ (labeled with the empty prefix). Completeness is immediate: if there is an SCE proof for $\Gamma \longrightarrow \Delta$, that very proof is also an SCL proof of $\Gamma \longrightarrow \Delta$. Conversely, the soundness theorem says that if there is an SCL proof of $\Gamma \longrightarrow \Delta$, then there is an SCE proof for $\Gamma \longrightarrow \Delta$. We establish this by simply adapting the general Herbrand theorem of [Lincoln and Shankar, 1994] to SCE. The idea behind the soundness result is that the structure of Herbrand terms provides enough information to reconfigure an SCL proof (by an inductive process of interchanges of inference, like that considered next in Section 2.4) so that equivalents of the eigenvariable conditions are enforced. The SCL proof may then be reindexed to respect SCE's eigenvariable requirements. ■

2.4 Permutability of inference and uniform proofs

Our syntactic methods for reasoning about derivations exploit *permutability of inference*—the general ability to transform derivations so that inferences are interchanged [Kleene, 1951]. To develop the notion of permutability of inference, we need to make some observations about the SCL sequent rules. First, the reasoning that is performed in subderivations is reasoning about subformulas (and vice versa). That is, in any spur sequent, the occurrence of the principal expression and the side expression all correspond to—or as we shall say, *are based in*—the occurrence of the principal expression in the base sequent. Likewise, each of the remaining expressions in the spur *are based in* an occurrence of an identical expression in the base. Here, as in [Kleene, 1951], we are assuming an *analysis* of each inference to specify this correspondence in the case where the same expression has multiple occurrences in the base or in a spur. Thus, our proof techniques, where they involve copying derivations, sometimes involve (implicit) reanalyses of inferences.

Now, in any derivation, the spur of one inference serves as the base for an *adjacent* inference or an axiom. We can therefore associate any tracked prefixed formula occurrence E in any sequent in the derivation with the occurrence in the root (or *end-sequent*) which E is based in. A similar notion can relate inferences, as follows. Suppose O is the inference at the root of a (sub)derivation, and L is another inference in the (sub)derivation. Then *L is based in* an expression E in the spur of O if the principal expression of L is based in E; *L is based in O* itself if E is a side expression of O. An important special case is that of an axiom based in an inference O. In effect, such an axiom marks a contribution that inference O contributes to completing the deduction.

To define interchanges of inference, we appeal to the two basic operations of *contraction* and *weakening*, which we cast as transformations on proofs. (In other proof systems, contraction and weakening may be introduced as explicit *structural rules*.)

Lemma 1 (Weakening) *Let $\mathcal{D}$ be an SCL proof, let Δ_0 be a finite multiset of tracked prefixed definitions and let Γ_0 be a finite multiset of tracked prefixed goals (in the same language as $\mathcal{D}$). Denote by $\Delta_0 + \mathcal{D} + \Gamma_0$ a derivation exactly like $\mathcal{D}$, except that where any node in $\mathcal{D}$ carries $\Delta \longrightarrow \Gamma$, the corresponding node in $\Delta_0 + \mathcal{D} + \Gamma_0$ carries $\Delta, \Delta_0 \longrightarrow \Gamma, \Gamma_0$. (When Δ_0 or Γ_0 is empty, we drop the corresponding $+$ from the notation.) Then $\Delta_0 + \mathcal{D} + \Gamma_0$ is also an SCL proof.*

Lemma 2 (Contraction) *Let $\mathcal{D}$ be an SCL proof whose root carries $\Delta \longrightarrow \Gamma, E, E$. Then we can construct an SCL proof $\mathcal{D}'$ whose root carries $\Delta \longrightarrow \Gamma, E$, whose height is at most the height of $\mathcal{D}$ and where there is a one-to-one correspondence (also preserving order of inferences) that takes any inference of $\mathcal{D}'$ to an inference with the same principal and side expressions in $\mathcal{D}$. We can likewise transform an SCL proof $\mathcal{D}$ whose root carries $\Delta, E, E \longrightarrow \Gamma$ into an SCL proof $\mathcal{D}'$ whose root carries $\Delta, E \longrightarrow \Gamma$.*

These lemmas follow from straightforward induction on the structure of derivations. These consequences continue to hold, suitably adapted, for the intermediate proof systems that we will construct from SCL in later sections.

Now consider two adjacent inferences in a derivation, a base inference R and an inference S (whose base is a spur of R). If S is not based in R, we may replace the derivation rooted at the base of R by a new derivation of the same end-sequent in which S applies at the root, R applies adjacent, and the remaining subderivations are copied from the original derivation (but possibly weakened to reflect the availability of additional logical premises). Performing such a replacement constitutes an interchange of rules R and S and demonstrates the permutability of R over S; see [Kleene, 1951]. SCL is formulated so that any such pair of inferences may be exchanged in this way.

We also observe that we can correctly introduce an abbreviation for goal occurrences of $\Box_i(A \supset B)$ by a single formula $(A >_i B)$ and the consolidation of corresponding inferences $(\rightarrow \Box_i)$ and $(\rightarrow\supset)$ into a single figure $(\rightarrow >_i)$, while retaining unrestricted interchange of inference. Again when the inference applies to principal A^{μ}_{X}, the figure is formulated using η for $\eta^{u}_{A}(\mu, X)$ as:

$$\frac{\Gamma, A^{\mu\eta}_{X,\mu\eta} \longrightarrow B^{\mu\eta}_{X,\mu\eta}, A >_i B^{\mu}_{X}, \Delta}{\Gamma \longrightarrow A >_i B^{\mu}_{X}, \Delta} \rightarrow >_i$$

We will refer to the calculus using $(\rightarrow >_i)$ in place of $(\rightarrow \Box_i)$ and $(\rightarrow \supset)$ as SCLI, and consider SCLI in the sequel.

[Miller, 1994, Miller, 1996] uses Definition 16 to characterize *abstract logic programming languages*.

Definition 16 *A cut-free sequent proof* $\mathcal{D}$ *is* uniform *if for for every subproof* $\mathcal{D}'$ *of* $\mathcal{D}$ *and for every non-atomic formula occurrence B in the right-hand side of the end-sequent of* $\mathcal{D}'$ *there is a proof* $\mathcal{D}''$ *that is equal to* $\mathcal{D}'$ *up to a permutation of inferences and is such that the base inference in* $\mathcal{D}''$ *introduces the top-level logical connective of B.*

Definition 17 *A logic with a sequent calculus proof system is an* abstract logic programming language *if restricting to uniform proofs does not lose completeness.*

It is easy to show that the sequent calculi SCL and SCLI are abstract logic programming languages in this sense. In fact, by this definition *every* SCL or SCLI derivation is uniform.

To anticipate our analysis of permutability in later sections, let us introduce the notion of an *eager* derivation in SCL or SCLI.

Definition 18 *Consider a derivation* $\mathcal{D}$ *containing a right inference R that applies to principal* E. *R is* delayed *exactly when there is a subderivation* $\mathcal{D}'$ *of* $\mathcal{D}$ *where:* $\mathcal{D}'$ *contains R;* $\mathcal{D}'$ *has a left inference L at the root; and the principal* E *of R is based in an occurrence of* E *in the end-sequent of* $\mathcal{D}'$.

Consider this schematic diagram of such a subderivation $\mathcal{D}'$:

$$\cfrac{\begin{array}{c}\cfrac{\vdots}{\ldots E \ldots}\,R \\ \downarrow\end{array}}{\ldots E \ldots}\,L$$

On an intuitive conception of a sequent proof as a record of proof search constructed from root upwards, R is delayed in that we have waited in $\mathcal{D}$ to apply R until after consulting the program by applying L, when we might have applied R earlier. Thus, we will also say in the circumstances of Definition 18 that R is delayed *with respect to L*.

Definition 19 $\mathcal{D}$ *is* eager *exactly when it contains no delayed applications of right rules.*

By transforming any derivation $\mathcal{D}$ into an eager derivation $\mathcal{D}'$ by permutations of inferences, we make it clear that reasoning about goals can always precede reasoning with program statements, and thereby provide a starting point for further analysis of goal-directed proof search.

Theorem 2 *Any SCL(I) derivation* $\mathcal{D}$ *is equal to an eager derivation* $\mathcal{D}'$ *up to permutations of inferences.*

The **proof** follows [Kleene, 1951, Theorem 2]. A double induction transforms each derivation into an eager one; the inner induction rectifies the final rule of a derivation whose subderivations are eager by an interchange of inferences (and induction) [Kleene, 1951, Lemma 10]; the outer one rectifies a derivation by rectifying the furthest violation from the root (and induction). See Appendix A. ■

3 Modular goal-directed proof search

3.1 Overview

Eager derivations do not make a satisfactory specification for goal-directed proof in a logic programming sense, because they do not embody a particularly directed search strategy. For one thing, eager derivations are free to work in parallel on different disjuncts of a goal using different program statements; in logic programming we want *segments* in which a single program statement and a single goal is in force. Moreover, eager derivations can reuse work across separate case analyses; in logic programming we want *blocks* where particular cases are investigated separately. Finally, because of their classical formulation, eager derivations do not enforce or exploit any modularity in their underlying logic. Our task is to remedy these faults of eager derivations.

Our result takes the form of an alternative sequent calculus SCLP which is equivalent to SCL. SCLP enforces a strictly goal-directed proof search through the structure of its inferences. First, SCLP sequents take the form

$$\Gamma; U \longrightarrow V; \Delta$$

We understand Γ to specify the global program and Δ to specify the global restart goals; both are multisets of tracked, prefixed formulas. U is at most one tracked, prefix formula, representing the current program statement; V is at most one tracked, prefixed formula, representing the current goal.

Logical rules apply only to the current program statement and the current goal. In addition, if there is a current program statement U then the current goal V must be an atomic formula. Thus, the interpreter first breaks the goal down into its components. Once an atomic goal is derived, the program is consulted; the selected program statement is decomposed and matched against the current goal by applicable logical rules. The form of the $(\supset\rightarrow)$ figure ensures that the interpreter continues to work on at most one goal at any time; this gives SCLP proofs their segment structure. Meanwhile, the form of the $(\vee \rightarrow)$ figures specify no current goal in its second case. The new current goal can then be chosen flexibly from possible restart goals. This gives SCLP proofs their block structure.

The new inferences are presented in Definition 20 and 21. Definition 20 shows the rules for decomposing program statements; Definition 21 shows the rules for decomposing goals.

Definition 20 (Logic programming calculus—programs) *The following inference figures describe the logic programming sequent calculus SCLP as it applies to program statements.*

1. *axiom—A atomic:*

$$\Gamma; A^{\nu} \longrightarrow A^{\nu}; \Delta$$

2. *decision (program consultation)—again A atomic:*

$$\frac{\Gamma, P^{\mu}_{X}; P^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta}{\Gamma, P^{\mu}_{X}; \longrightarrow A^{\nu}_{Y}; \Delta}\ \textit{decide}$$

3. *conjunctive:*

$$\frac{\Gamma; P^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta}{\Gamma; P \wedge Q^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta}\ \wedge\rightarrow_{L}$$

$$\frac{\Gamma; Q^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta}{\Gamma; P \wedge Q^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta} \wedge \rightarrow_R$$

4. *disjunctive:*

$$\frac{\Gamma; P^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta \qquad \Gamma, Q^{\mu}_{X}; \longrightarrow; \Delta}{\Gamma; P \vee Q^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta} \vee \rightarrow_L$$

$$\frac{\Gamma; Q^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta \qquad \Gamma, P^{\mu}_{X}; \longrightarrow; \Delta}{\Gamma; P \vee Q^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta} \vee \rightarrow_R$$

5. *implication:*

$$\frac{\Gamma; \longrightarrow Q^{\mu}_{X}; \Delta \qquad \Gamma; P^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta}{\Gamma; Q \supset P^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta} \supset\rightarrow$$

6. *necessity—subject to the side condition that there is a typing derivation $S, \Xi_{\nu} \triangleright \mu/\mu\nu : i$:*

$$\frac{\Gamma; P^{\mu\nu}_{X,\mu\nu} \longrightarrow A^{\nu'}_{Y}; \Delta}{\Gamma, \Box_i P^{\mu}_{X} \longrightarrow A^{\nu'}_{Y}; \Delta} \Box_i \rightarrow$$

7. *existential—subject to the side condition that h is $h^{u}_{\exists xP}(\mu, X)$ for some u:*

$$\frac{\Gamma; P[h/x])^{\mu}_{X,h} \longrightarrow A^{\nu}_{Y}; \Delta}{\Gamma; \exists x.P^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta} \exists \rightarrow$$

8. *universal—subject to the side condition that there is a typing derivation $S, \Xi_{t,\mu} \triangleright t : \mu$:*

$$\frac{\Gamma; P[t/x]^{\mu}_{X,t} \longrightarrow A^{\nu}_{Y}; \Delta}{\Gamma; \forall x.P^{\mu}_{X} \longrightarrow A^{\nu}_{Y}; \Delta} \forall \rightarrow$$

Definition 21 (Logic programming calculus—goals) *The following inference figures describe the logic programming sequent calculus SCLP as it applies to goals.*

1. *restart:*

$$\frac{\Gamma; \longrightarrow G^{\nu}_{X}; G^{\nu}_{X}, \Delta}{\Gamma; \longrightarrow; G^{\nu}_{X}, \Delta} \textit{restart}$$

2. *conjunctive goals:*

$$\frac{\Gamma; \longrightarrow F^{\mu}_{X}; \Delta \qquad \Gamma; \longrightarrow G^{\mu}_{X}; \Delta}{\Gamma; \longrightarrow F \wedge G^{\mu}_{X}; \Delta} \rightarrow \wedge$$

3. *disjunctive goals:*

$$\frac{\Gamma; \longrightarrow F^{\mu}_{X}; \Delta}{\Gamma; \longrightarrow F \vee G^{\mu}_{X}; \Delta} \rightarrow \vee_L$$

$$\frac{\Gamma; \longrightarrow G^{\mu}_{X}; \Delta}{\Gamma; \longrightarrow F \vee G^{\mu}_{X}; \Delta} \rightarrow \vee_R$$

4. *necessary goals—where η is $\eta^u_A(\mu, X)$ for A^μ_X the principal of the rule and for some u for which η^u_A does not occur in Δ or Γ:*

$$\frac{\Gamma, F^{\mu\eta}_{X,\mu\eta}; \longrightarrow G^{\mu\eta}_{X,\mu\eta}; G^{\mu\eta}_{X,\mu\eta}, \Delta}{\Gamma; \longrightarrow F >_i G^\mu_X; \Delta} \rightarrow \Box_i \supset$$

$$\frac{\Gamma; \longrightarrow G^{\mu\eta}_{X,\eta}; G^{\mu\eta}_{X,\mu\eta}, \Delta}{\Gamma; \longrightarrow \Box_i G^\mu_X; \Delta} \rightarrow \Box_i$$

5. *universal goals—subject to the side condition that h is $h^u_{\forall x G}(\mu, X)$ for some u:*

$$\frac{\Gamma; \longrightarrow G[h/x]^\mu_{X,h}; \Delta}{\Gamma; \longrightarrow \forall x. G^\mu_X; \Delta} \rightarrow \forall$$

6. *existential goals—subject to the side condition that there is a typing derivation $\mathcal{S}, \Xi_{t,\mu} \triangleright t : \mu$:*

$$\frac{\Gamma; \longrightarrow G[t/x]^\mu_{X,t}; \Delta}{\Gamma; \longrightarrow \exists x. G^\mu_X; \Delta} \rightarrow \exists$$

Inspection of the figures of Definitions 20 and 21 reveals the following generalization of modularity and locality: in any derivation, the label of the current program statement must be a prefix of the label of the current goal. Moreover, goal labels are always extended with novel symbols, because of the eigenvariable condition in the $(\rightarrow \Box)$ figure. Inductively, these facts determine a strong invariant—consider a block beginning with a restart inference whose spur is

$$\Gamma; \longrightarrow G^\nu_X; \Delta$$

and consider any expression P^μ_Y in Γ. If μ is not a prefix of ν, then μ will not be a prefix of the label of any goal formula in the block. Thus P^μ_Y cannot be used in the block. (Compare [Stone, 1999, Lemma 2].)

This is why the (restart) rule of SCLP can be made modular, so that it limits the work that is reanalyzed to the scope of the ambiguity just introduced. We must simply show that the new disjunct will contribute to its restart goal. In particular, define canceled blocks as in Definition 23.

Definition 22 (Linked) *An expression E in a sequent in an SCLU derivation $\mathcal{D}$ is* linked *if the principal formula of an axiom in the same block of $\mathcal{D}$ as that sequent is based in E. An inference R is* linked *in $\mathcal{D}$ if some side expression of R is linked in each spur of R. A derivation or block is* linked *iff all of the inferences in it are linked.*

Definition 23 (Canceled) *A block is* canceled *if it contains the root of $\mathcal{D}$, or if the side expression E of the $(\vee \rightarrow)$ inference whose spur is the root of the block is linked.*

Thus a canceled block includes a use of any disjunctive case introduced in the block. The key fact about SCLP is that it suffices to consider only canceled blocks in proof search.

$$\boxed{2'}\quad \dfrac{\dfrac{\dfrac{\dfrac{\ldots;C \longrightarrow C;F}{\ldots,C; \longrightarrow C;F}\,(\text{decide}) \qquad \ldots;F \longrightarrow F;F}{C,\ldots;C\supset F \longrightarrow F;F}\,(\supset\rightarrow)}{C\supset F,C\ldots; \longrightarrow F;F}\,(\text{decide})}{A\vee B,C\vee D,A\supset F,C\supset F,(B\wedge D)\supset F,B,C; \longrightarrow ;F}\,(\text{restart})$$

$$\boxed{3'}\quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\ldots;B \longrightarrow B;F}{B,\ldots; \longrightarrow B;F}\,(\text{decide}) \qquad \dfrac{\dfrac{\ldots;D \longrightarrow D;F \qquad \boxed{2'}}{\ldots;C\vee D \longrightarrow D;F}\,(\vee\rightarrow_R)}{C\vee D,\ldots; \longrightarrow D;F}\,(\text{decide})}{C\vee D,B,\ldots; \longrightarrow B\wedge D;F}\,(\rightarrow\wedge) \qquad \ldots;F \longrightarrow F;F}{C\vee D,B,\ldots;(B\wedge D)\supset F \longrightarrow F;F}\,(\supset\rightarrow)}{C\vee D,(B\wedge D)\supset F,B,\ldots; \longrightarrow F;F}\,(\text{decide})}{A\vee B,C\vee D,A\supset F,C\supset F,(B\wedge D)\supset F,B; \longrightarrow ;F}\,(\text{restart})$$

$$\boxed{1}\quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\ldots;A \longrightarrow A;F \qquad \boxed{3'}}{\ldots;A\vee B \longrightarrow A;F}\,(\vee\rightarrow_L)}{A\vee B,\ldots; \longrightarrow A;F}\,(\text{decide}) \qquad \ldots;F \longrightarrow F;F}{A\vee B,\ldots;A\supset F \longrightarrow F;F}\,(\supset\rightarrow)}{A\vee B,A\supset F,\ldots; \longrightarrow F;F}\,(\text{decide})}{A\vee B,C\vee D,A\supset F,C\supset F,(B\wedge D)\supset F; \longrightarrow ;F}\,(\text{restart})$$

Figure 6: The SCLP presentation of the proof of Figure 5.

Theorem 3 *Let Γ and Δ be multisets of tracked prefixed expression in which each formula is tracked by the empty set and prefixed by the empty prefix. There is a proof of $\Gamma \longrightarrow \Delta$ in SCL exactly when there is a proof of $\Gamma; \longrightarrow ;\Delta$ in SCLP in which every block is canceled.*

The discussion of the following subsections represents an outline of the proof of this result. The strategy is to transform eager proofs from SCL to SCLP by a series of refinements of sequent rules that make the logic programming strategy explicit. We give force to the idea that the interpreter has a current goal and current program statement, in Section 3.2. Then we create blocks for case analysis, in Section 3.3. Finally, we enforce modularity, in Section 3.4. See also Appendix B.

Figure 6 shows how the proof of Figure 5 is recast in SCLP. Figure 6 extends Figure 5 to make the bookkeeping of goal-directed proof explicit. In Figure 6, the informal underline of Figure 5 is gone, and instead the current goal and the current program statement are displayed at distinguished positions in sequents. New (restart) and (decide) inferences mark the consideration of new goals and new program statements. Of course, the logical content of the two inferences is identical. Applying Definition 23, block $\boxed{1}$ is canceled because it contains the root; there is no new disjunct

$$\boxed{5}\quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\ldots;B^{\alpha} \longrightarrow B^{\alpha};\ldots}{\ldots,B^{\alpha}; \longrightarrow B^{\alpha};\ldots}\,(\text{decide}) \qquad \ldots;A^{\alpha} \longrightarrow A^{\alpha};\ldots}{B^{\alpha},\ldots;B \supset A^{\alpha} \longrightarrow A^{\alpha};\ldots}\,(\supset\to)}{B^{\alpha},\ldots;\Box(B \supset A) \longrightarrow A^{\alpha};\ldots}\,(\Box\to)}{\Box(B \supset A),B^{\alpha}; \longrightarrow A^{\alpha};\ldots}\,(\text{decide})}{\Box(A \vee B),\Box(B \supset A),\Box(C \vee D),\Box(D \supset C),B^{\alpha}; \longrightarrow ;A^{\alpha},(\Box A) \wedge (\Box C)}\,(\text{restart})$$

$$\boxed{6}\quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\ldots;D^{\beta} \longrightarrow D^{\beta};\ldots}{\ldots,D^{\beta}; \longrightarrow D^{\beta};\ldots}\,(\text{decide}) \qquad \ldots;C^{\beta} \longrightarrow C^{\beta};\ldots}{D^{\beta},\ldots;D \supset C^{\beta} \longrightarrow C^{\beta};\ldots}\,(\supset\to)}{D^{\beta},\ldots;\Box(D \supset C) \longrightarrow C^{\beta};\ldots}\,(\Box\to)}{\Box(D \supset C),D^{\beta}; \longrightarrow C^{\beta};\ldots}\,(\text{decide})}{\Box(A \vee B),\Box(B \supset A),\Box(C \vee D),\Box(D \supset C),D^{\beta}; \longrightarrow ;C^{\beta},(\Box A) \wedge (\Box C)}\,(\text{restart})$$

$$\boxed{4}\quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\ldots;A^{\alpha} \longrightarrow A^{\alpha};\ldots \qquad \boxed{5}}{\ldots;A \vee B^{\alpha} \longrightarrow A^{\alpha};\ldots}\,(\vee\to_L)}{\ldots;\Box(A \vee B) \longrightarrow A^{\alpha};\ldots}\,(\Box\to)}{\Box(A \vee B),\ldots; \longrightarrow A^{\alpha};\ldots}\,(\text{decide})}{\Box(A \vee B),\ldots; \longrightarrow \Box A;\ldots}\,(\to\Box) \qquad \dfrac{\dfrac{\dfrac{\dfrac{\ldots;C^{\beta} \longrightarrow C^{\beta};\ldots \qquad \boxed{6}}{\ldots;C \vee D^{\beta} \longrightarrow C^{\beta};\ldots}\,(\vee\to_L)}{\ldots;\Box(C \vee D) \longrightarrow C^{\beta};\ldots}\,(\Box\to)}{\Box(C \vee D),\ldots; \longrightarrow C^{\beta};\ldots}\,(\text{decide})}{\Box(C \vee D),\ldots; \longrightarrow \Box C;\ldots}\,(\to\Box)}{\Box(A \vee B),\Box(C \vee D),\ldots; \longrightarrow (\Box A) \wedge (\Box C);\ldots}\,(\to\wedge)}{\Box(A \vee B),\Box(B \supset A),\Box(C \vee D),\Box(D \supset C); \longrightarrow ;(\Box A) \wedge (\Box C)}\,(\text{restart})$$

Figure 7: The SCLP presentation of the proof of Figure 2. We suppress tracking of formulas and hide the internal structure of Herbrand terms.

to discharge here. Block $\boxed{3'}$ is canceled: the inference whose spur is the root of block $\boxed{3'}$ is the $(\vee \to_L)$ and its side expression is an occurrence of B, the new disjunct in the block. This occurrence is linked in the block because of the leftmost axiom $\ldots;B \longrightarrow B;F$ which is based in it; the inference $(\vee \to_L)$ is linked in the block for the same reason. Similarly block $\boxed{2'}$ is canceled because the new disjunct C (the side expression of the $(\vee \to_R)$ inference whose spur is the root of block $\boxed{2'}$) contributes to the leftmost axiom $\ldots;C \longrightarrow C;F$ in the block.

Figure 7 shows how the proof of Figure 2 is recast in SCLP. The most dramatic change here is that the inferences of Figure 7 are segmented out into three blocks. Another change is the discipline of explicit scope; we introduce a suitable term α to represent the generic context in which we prove $\Box A$ and another suitable term β to represent the generic context in which we prove $\Box C$. Correspondingly, we transition to α in using $\Box(A \vee B)$ and transition to β in using $\Box(C \vee D)$. In the (restarts) of $\boxed{5}$ and $\boxed{6}$ the changes interact. In $\boxed{5}$ we pick the modular restart A^{α} in order to permit a contribution by the new assumption B^{α}. In $\boxed{6}$ we pick the modular restart C^{β} in order to permit a contribution by the new assumption D^{β}.

3.2 Segment structure

Our first task is to formalize goal-directed search that directs attention to a single goal at a time. To distinguish such goals, we begin with a trick that for now is purely formal—introducing an *articulated* SCLI. We represent assumptions as a pair $\Pi;\Gamma$ with Π encoding the global program and Γ encoding local program statements; eventually local statements will be processed only in the current segment and then discarded. (Compare the similar notation and treatment from [Girard, 1993].) Similarly, we represent goals as a pair $\Delta;\Theta$, with Θ encoding the restart goals and Δ encoding the local goals; ultimately, we will also describe inference rules which will discard Δ between segments. With this representation, principal formulas of logical rules are local formulas, in Γ or Δ; so are the side formulas—with these exceptions: the $(\rightarrow \Box)$ and $(\rightarrow >)$ rules augment Π instead of Γ (when they add a new program statement) and Θ instead of Δ (when they add new restart goals).

New (decide) and (restart) rules keep this change general; they allow a global formula—a program statement or restart goal—to be selected and added to the local state.

$$\frac{\Pi, A^{\mu}_{X};\Gamma, A^{\mu}_{X} \longrightarrow \Delta;\Theta}{\Pi, A^{\mu}_{X};\Gamma \longrightarrow \Delta;\Theta}\,\text{(decide)} \qquad \frac{\Pi;\Gamma \longrightarrow \Delta, G^{\mu}_{X};\Theta, G^{\mu}_{X}}{\Pi;\Gamma \longrightarrow \Delta;\Theta, G^{\mu}_{X}}\,\text{(restart)}$$

Lemma 3 (Articulation) *Every SCLI deduction can be converted into an articulated SCLI deduction with an end-sequent of the form* $\Pi; \longrightarrow;\Theta$ *in such a way that if the initial derivation is eager then so is the resulting derivation (and vice versa).*

Proof. Straightforward structural induction. ■

The next step is to introduce an inference figure $(\supset\rightarrow^{S})$ that imposes a *segment* structure on derivations, thus:

$$\frac{\Pi; \longrightarrow A^{\mu}_{X},\Delta;\Theta \qquad \Pi;\Gamma, A\supset B^{\mu}_{X}, B^{\mu}_{X} \longrightarrow \Delta;\Theta}{\Pi;\Gamma, A\supset B^{\mu}_{X} \longrightarrow \Delta;\Theta}\,(\supset\rightarrow^{S})$$

Definition 24 (Segment) *A* segment *in a derivation* $\mathcal{D}$ *is a maximal tree of contiguous inferences in which the left subtree of any* $(\supset\rightarrow^{S})$ *inference is omitted.*

The distinctive feature of the $(\supset\rightarrow^{S})$ figure is that the local results inferred from the program are discarded in the subderivation where the new goal is introduced. In an eager derivation, this will begin a new segment where first the new goal will be considered and then a new program statement will be selected to establish that goal.

We will define two calculi using $(\supset\rightarrow^{S})$. The first, SCLS, eliminates the $(\supset\rightarrow)$ inference of the articulated SCLI and instead has $(\supset\rightarrow^{S})$. The second, SCLV, is a calculus like the articulated SCLI but also allows $(\supset\rightarrow^{S})$; $(\supset\rightarrow)$ and $(\supset\rightarrow^{S})$ can appear anywhere in an SCLV derivation. We introduce SCLV to facilitate the incremental transformation of articulated SCLI proofs into SCLS proofs.

Lemma 4 *An eager articulated SCLI derivation whose end-sequent is of the form*

$$\Pi; \rightarrow \Delta;\Theta$$

can be transformed to an eager SCLS derivation of the same end-sequent.

Proof. We proceed with an inductive construction that eliminates $(\supset\to)$ inferences in favor of $(\supset\to^S)$ inferences one at a time. See Appendix B.1. ∎

3.3 Block structure

We now revise how we perform case analysis from assumptions. We introduce new rules where all local work is discarded in the subderivation written on the right. This corresponds to a sequent of the form $\Pi; \longrightarrow; \Theta$. In addition, some *global* work may be discarded in the right subderivation; this helps clarify the structure of derivations. Accordingly, there may be additional formula occurrences Π' and Θ' in the base sequent that are not copied up to the right subderivation. Finally, the right subderivation may address either the (textually) first disjunct or the second disjunct. This leads to the two inference figures below.

$$\frac{\Pi,\Pi';\Gamma,A\vee B^{\mu},A^{\mu} \longrightarrow \Delta;\Theta,\Theta' \qquad \Pi,B^{\mu};\longrightarrow;\Theta}{\Pi,\Pi';\Gamma,A\vee B^{\mu} \longrightarrow \Delta;\Theta,\Theta'}\vee\to^B_L$$
$$\frac{\Pi,\Pi';\Gamma,A\vee B^{\mu},B^{\mu} \longrightarrow \Delta;\Theta,\Theta' \qquad \Pi,A^{\mu};\longrightarrow;\Theta}{\Pi,\Pi';\Gamma,A\vee B^{\mu} \longrightarrow \Delta;\Theta,\Theta'}\vee\to^B_R$$

We call these inferences *blocking* $(\vee\to)$ inferences, or $(\vee\to^B)$ inferences. We will appeal to two calculi in which these inferences appear. The first, SCLU, permits both ordinary $(\vee\to)$ and $(\vee\to^B)$ inferences, without restriction. SCLU is convenient for describing transformations between proofs. The second, SCLB, permits $(\vee\to^B)$ inferences but not ordinary $(\vee\to)$ inferences.

Blocks are more than just boundaries in the proof; they provide a locus for enforcing modularity. We will ensure that a disjunct contributes inferences to the new block where it is introduced. Thanks to this contribution, we can narrow down the choice of goals to restart in a modular way.

This result is made possible only by maintaining the right structure as we introduce $(\vee\to^B)$ inferences. We use path prefixes to make explicit connections between program statements and any goals that they help establish. The key notions are *spanning*, *simplicity* and *balance* for sequents. Spanned, simple and balanced sequents represent a consistent evolution of the state of proof search, which records a full set of restart goals and the corresponding assumptions, with no redundancy.

Definition 25 (Carrier) *The* carrier *of a non-empty Herbrand prefix $\mu\eta$ is $B^{\mu\eta}_{X,\mu\eta}$ if η is $\eta^u_{A>_iB}(\mu,X)$ and otherwise, when η is $\eta^u_{\Box_iA}(\mu,X)$, is $A^{\mu\eta}_{X,\mu\eta}$.*

Definition 26 (Spanned) *Say one multiset of tracked prefixed formulas, Π, is* spanned *by another, Θ, if for every expression occurrence A^{μ}_{X} in Π and every non-empty prefix ν of μ there is an occurrence of the carrier of ν in Θ. It is easy to see there is a minimal set Θ that spans Π and that such Θ spans itself. A sequent $\Pi;\Gamma \longrightarrow \Delta;\Theta$ is* spanned *if Π is spanned by Θ, Γ is spanned by Θ, Δ is spanned by Θ and Θ is spanned by Θ. A derivation or block is* spanned *if every sequent in it is spanned.*

Definition 27 (Simple) *A multiset Ψ is* simple *if no expression occurs multiple times in Ψ; a sequent of the form $\Pi;\Gamma \longrightarrow \Delta;\Theta$ is* simple *if Π and Θ are simple. A derivation or block is* simple *iff every sequent in it is simple.*

Definition 28 (Balanced) *A pair of multisets of tracked, prefixed formulas* Π, Θ *is* balanced *if*

- *for any* $\eta = \eta^u_{B>_iC}(\mu, X)$, η *occurs in* Θ *exactly when the expression* $B^{\mu\eta}_{X,\mu\eta}$ *occurs in* Π *and exactly when the expression* $C^{\mu\eta}_{X,\mu\eta}$ *occurs in* Θ*; and*
- *for any* $\eta = \eta^u_{\Box A}(\mu, X)$, η *occurs in* Θ *exactly when the expression* $A^{\mu}_{X,\mu\eta}$ *occurs in* Θ.

A sequent $\Pi; \Gamma \longrightarrow \Delta; \Theta$ *is* balanced *if the pair* Π, Θ *is balanced. A block or derivation is* balanced *if every sequent in the block is balanced.*

We use the notion of an *isolated block* to obtain an even stronger characterization of proof search that proceeds in a well-regimented way. In an isolated block, the only expressions preserved across a blocking inference are those that are in some sense intrinsic to the restart problem created by that inference. Specifically, each nested block must begin with the same end-sequent as the outer block, except for additional program statements that have to be added in order to introduce the newly-assumed disjunct, and the further goal and program statements required to obtain a balanced and spanned sequent.

Definition 29 (Isolated) *Let* $\mathcal{D}$ *be an SCLU derivation, and let* $\mathcal{B}$ *be a block of* $\mathcal{D}$*. Write the end-sequent of* $\mathcal{B}$ *as* $\Pi; \Gamma \longrightarrow \Delta; \Theta$ *and consider the right subproof of some* $(\vee \rightarrow^B)$ *inference L at the boundary of* $\mathcal{B}$ *has an end-sequent of the form* $\Pi', E; \longrightarrow; \Theta'$*. The* exported *expressions in* Π', Π'_e, *consist of the occurrences of expressions F in* Π' *such that either is F based in an occurrence of F in* Π *or is based in an occurrence of F as the side expression of an inference in which E is also based.*

$\mathcal{B}$ *is* isolated *if the right subproof of each* $(\vee \rightarrow^B)$ *inference L at the boundary of* $\mathcal{B}$ *has an end-sequent of the form* $\Pi', E; \longrightarrow; \Theta'$ *meeting the following conditions: E is the side-expression of L;* Θ' *is the minimal multiset of expressions which spans* Π'_e, E, Θ *and includes* Θ*; and* Π' *is the smallest multiset including* Π'_e, E *for which* Π', Θ' *is balanced.* $\mathcal{D}$ *is* isolated *iff every block of* $\mathcal{D}$ *is isolated.*

Isolation allows us to keep close tabs on the uses of formulas within blocks, which is important for establishing modularity later. In particular, isolation provides a key notion in formalizing the obvious fact that an inference that makes no contribution to an SCLU derivation can be omitted.

Finally at this stage, we refine the form of proofs which we are willing to count as goal-directed. Now it will often happen that, while each block of a derivation may be eager, the derivation as a whole will not be eager. As observed in [Nadathur and Loveland, 1995], derivations with blocks can nevertheless be seen as eager throughout by reconstructing the (restart) rule as backchaining against the negation of a subgoal. But we will simply consider *blockwise eager* derivations from now on.

Definition 30 (Blockwise delayed) *R is* blockwise delayed *exactly when there is a tree of contiguous inferences* $\mathcal{D}'$ *within a single block of* $\mathcal{D}$ *where:* $\mathcal{D}'$ *contains R;* $\mathcal{D}'$ *has a left inference L at the root; and the principal E of R is based in an occurrence of E in the end-sequent of* $\mathcal{D}'$.

Definition 31 (Blockwise eager) *D is* blockwise eager *exactly when it contains no blockwise delayed applications of right rules.*

Obviously, we can use weakening to transform an SCLB or SCLU derivation into a SCLS derivation, so the blocking inference figures are sound. The completeness of SCLB is a consequence of Lemma 5.

Lemma 5 *We are given a blockwise eager SCLS derivation $\mathcal{D}$ whose end-sequent is spanned and balanced and takes the form:*

$$\Pi; \longrightarrow; \Theta$$

We transform $\mathcal{D}$ into a blockwise eager SCLB derivation in which every block is canceled, linked, isolated, simple, balanced and spanned.

Proof. We can transform individual blocks to achieve a streamlined form, which already implicitly reflects the logic programming search strategy of focused search on particular goals and program statements. By pursuing a suitable ordering strategy as we inductively repeat this inductive transformation, we can create the desired SCLB proofs with an overall modular block structure. See Appendix B.2. ■

3.4 Modularity

We now derive SCLP from SCLB. SCLP proofs can be rewritten to SCLB rules by a weakening transformation. Conversely, rewriting SCLB proofs to SCLP proofs is accomplished by induction on the structure of proofs. The transformation is possible because multiple formulas in sequents are needed only for passing ambiguities and work done across branches in the search; this is ruled out by the use of $(\vee \rightarrow^B_L)$, $(\vee \rightarrow^B_R)$ and $(\supset\rightarrow^S)$.

Lemma 6 *Given a blockwise eager SCLB derivation $\mathcal{D}$, with end-sequent*

$$\Pi; \longrightarrow; \Theta$$

in which every block is linked, simple and spanned, we can construct a corresponding SCLP derivation of the same end-sequent in which every block remains linked.

Proof. By induction on the structure of proofs. See Appendix B.3. ■

4 Assessment and conclusions

To execute modal specifications requires leveraging both the flexibility of efficient classical theorem-proving and the distinctive modularity of modal logic. This is a significant problem because the two are at odds. On the one hand, flexible search strategies impose no constraints on the relationships among inferences. By ignoring modularity, they can leave open inappropriate possibilities for search. On the other hand, brute-force modular systems may place such strong constraints on the order in which search must proceed that it becomes impossible to guide that search in a predictable, goal-directed way. In this paper, we have explored one strategy for balancing the flexibility of classical goal-directed search with the modularity of modal logic. This strategy culminates in the development of a modular logic programming sequent calculus SCLP.

[Stone, 1998b] describes a preliminary implementation of proof search in SCLP as a logic programming interpreter DIALUP. I close by summarizing *how* (Section 4.1) and no less importantly *why* (Section 4.2) I developed this implementation.

4.1 Implementation

An effective implementation of SCLP requires further treatments of *unification* and *search control*.

In general, to implement first-order sequent calculus proof search, we must *lift* the inference figures. That is, we adapt the inferences that require instantiation to specific terms so that they introduce *logic variables* instead. As we construct the proof, we accumulate *constraints* on the values of these variables—for example, we get constraints when an axiom link in the proof requires two formulas to be identical. In the lifted system, each proof we find represents the set of ground proofs that you get by replacing the variables with values that satisfy the constraints. Lifting is the essence of the resolution procedure [Robinson, 1965] but can be regarded as a general metatheoretical strategy. [Lincoln and Shankar, 1994, Voronkov, 1996] offer particularly general discussions of this strategy at its most sophisticated.

For first-order modal inference in prefixed calculi, lifting introduces two kinds of logic variables, and two corresponding kinds of constraints. First-order quantifiers introduce logic variables over individuals, subject to the familiar constraints that give rise to term unification problems. Modal inferences, meanwhile, introduce logic variables over prefixes, subject to path equations. This leads to specialized problems of equational unification; good solutions are known for the general setting of multi-modal logic; see for example [Auffray and Enjalbert, 1992, Debart et al., 1992, Otten and Kreitz, 1996, Schmidt, 1998].

The logical fragment of SCLP makes path equations particularly simple. Inspection of the SCLP proof rules shows that, at any point in proof search, we have enough path constraints to determine *ground* substitutions for all the path variables in the sequent except possibly for variables in the current program statement that are about to be unified with a goal. In many cases, this makes path equations easy to solve—a compact representation of all possible solutions can be computed in polynomial time. The details are beyond the scope of this paper, but see [Stone, 1998b].

Search control is the other issue. An implementation has to make commitments about what statements to try and what rules to use to process those statements. The fact that SCLP program and goal statements are labeled with ground prefixes means that we can easily test that a statement's label is a prefix of the goal label before attempting to match the statement and the goal. We can also identify an atomic subformula of the statement nondeterministically as the *head*, and commit to match that head with the goal. Before doing so, we can for example test that the head and the goal share the same predicate symbol.

In the case of disjunction, we also want to make sure that we avoid reporting duplicate proofs, despite the duplicate rules for disjunction that we have. Loveland considers a number of heuristics for this [Loveland, 1991], and we expect that they apply in SCLP as well as in Near-Horn Prolog. But here is another heuristic. As motivated in Section 1.1.3, $(\vee \rightarrow_R)$ is required only for cancellation. When we use it, we expect to cancel an assumption (like B in Figure 5) that could not be canceled otherwise. We can make this precise: $(\vee \rightarrow_R)$ should only be used in a restart block, and the assumption that is canceled in that block ought not to be used in the subsequent restart block initiated by the $(\vee \rightarrow_R)$ inference. Otherwise, we will independently construct an alternative proof that uses $(\vee \rightarrow_L)$ instead. Naturally, the kind of block analysis illustrated in the proof of Theorem 3 can be used to show that this restriction is complete.

4.2 Applications in modal representation

In classical logic, indefinite information is a bit exotic. Rather than developing an indefinite specification, we much prefer to collect the additional information required to describe the world in a precise, definite way. This is not true at all with modal specifications. Modal specifications get much of their interest from their ability to contrast different perspectives or sources of information. What one source of information represents with specific, definite information, another source represents with abstract, indefinite information. Computation from modal specifications involves the coordinated exchange of information between these sources.

In particular, problems of *planning* [Stone, 1998a] and problems of *communication* [Stone, 2000] depend on indefinite modal specifications. In planning, one agent, the *scheduler*, has to allocate a task to another agent, the *executive*. (The executive may just be the scheduler at a later point in time!) It is unrealistic to expect that the scheduler will know *exactly* what the executive *will* do; this almost certainly requires information that is not available to the scheduler. Rather, the scheduler should merely know what the executive *can* do. This means that, to be useful, the scheduler must have an *indefinite* modal specification that abstractly describes the information that will be available to the executive. For examples, see [Moore, 1985, Morgenstern, 1987, Scherl and Levesque, 1993, Davis, 1994] as well as [Stone, 1998a].

In communication, the task of one agent, the *speaker*, is to formulate an utterance that allows another agent, the *hearer*, to answer a question. There are many cases where the speaker does not have enough information to answer the question directly. However, the speaker can still design an utterance that allows the hearer to infer the right answer, because the hearer knows something the speaker does not. Concretely, a user of a computer interface might want to know what action to take next. The right answer might be for the user to type *jdoe* into a certain text box. The speaker might know to say *enter your user ID*, even if the speaker does not know what the user ID is. Again, the speaker can make such choices meaningfully only from an indefinite modal specification that says what the hearer knows abstractly but not definitely. See [Stone, 2000] for a worked-out formal case study.

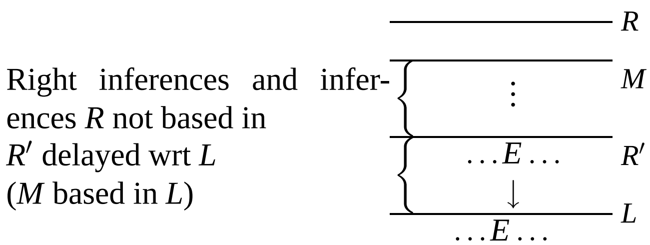
A Proof of Theorem 2

Any SCL(I) derivation $\mathcal{D}$ is equal to an eager derivation $\mathcal{D}'$ up to permutations of inferences.

The proof depends on a generalization of delayed inferences, which we can term *misplaced* inferences since we intend to eliminate them. We assume an overall derivation $\mathcal{D}$, and consider a right inference R that applies to principal E within some subderivation $\mathcal{D}'$ of $\mathcal{D}$.

Definition 32 *We say a right inference R is* right-based *on an inference R' in $\mathcal{D}$ if $R = R'$ or R is based on R' and every inference on which R is based above and including R' is a right inference. Then R is* misplaced *in $\mathcal{D}'$ exactly when there are inferences M and R' in $\mathcal{D}'$ such that, in $\mathcal{D}$, M is based on an inference L, R is right-based on R', and R' is delayed with respect to L.*

In this case we will also say R is misplaced *with respect to M*. We can abstract a key case of misplaced inferences by the following schematic derivation:



This schematic derivation shows informally how *misplaced inferences* help provide an inductive characterization of the inferences that stand in the way of obtaining an eager derivation. In an eager derivation, it will be impossible for R to appear above L. For R' cannot be delayed with respect to L, but once R' and L are interchanged, we will obtain a new delayed inference that R is based in, until finally we must interchange L and R. Of course, to do this, we must first interchange R with the *misplaced* inferences, such as M, which stand between R and L and cannot themselves be interchanged with L because they are based in L.

Observe that the relation R is misplaced with respect to M is asymmetrical. To see this, suppose R is misplaced with respect to M. By definition, R is right-based on R' which is delayed with respect to a left inference L on which M is based. Meanwhile, for M to be misplaced with respect to R, by definition, we must have M right-based on M' and R based in some left rule L_R. Any such M' would have to be based in L since no left inferences intervene between M and M'; M' must thus appear *inside* a schematic like that above. At the same time, since no left inferences intervene between R and R', R' would have to be based in any such L_R, which must thus appear *outside* such a schematic, closer to the root of the overall derivation. Accordingly, any such L_R must occur closer to the root of $\mathcal{D}$ than L; meanwhile the principal of M' is introduced further from the root than L. So we will not have M' delayed with respect to L_R.

Call R *badly misplaced* in $\mathcal{D}'$ if R is misplaced with respect to M and M occurs closer to the root than R. A subderivation $\mathcal{D}'$ with no badly misplaced inferences will be called *good*. An overall good derivation is also eager, since any delayed inference is badly misplaced.

We can now present the proof in full using a lemma.

Lemma 7 *Consider a subderivation $\mathcal{D}'$ of an overall derivation $\mathcal{D}$, with the property that $\mathcal{D}'$ has good immediate subderivations and that $\mathcal{D}'$ ends in inference M. From $\mathcal{D}'$ we can construct a derivation with the same end-sequent that is good.*

Proof. The assumption that the immediate subderivations of $\mathcal{D}'$ are good is a very powerful one. For suppose that some inference is badly misplaced with respect to some other in $\mathcal{D}'$. Then we can only have some rule R badly misplaced with respect to M—anything else would contradict that assumption.

In fact, we can show that some such R must be adjacent to M. Consider an inference S that intervenes between R and M: we will show that S must be badly misplaced with respect to M too. By the definition of misplaced, M is based on some left rule L in $\mathcal{D}$, R is right-based on R', and R' is delayed with respect to L. Now consider the inferences that S is based on above L. If any of these is a left inference L', or S is itself a left inference, then R is also misplaced with respect to S—indeed, badly misplaced. This contradicts the assumption that the subderivations of $\mathcal{D}'$ are good. So none of these inferences can be a left inference, which means S is a right inference that is right-based on some inference S' above L. S' must be delayed with respect to L. Hence S is badly misplaced with respect to M.

Now we can proceed after [Kleene, 1951, Lemma 10]. Define the *grade* of D' as the number of badly misplaced inferences in D'. We show by induction on the grade that D' can be transformed to a good one.

The base case is a derivation of grade 0. This case has D' itself good. Thus, suppose the lemma holds for derivations of grade g, and consider D' of grade $g+1$. By the argument just given, one immediate subderivation—call it D''—must end with an inference R which is badly misplaced with respect to M. Such an R of course cannot be based in M, so we interchange inferences R and M. In the result, the subderivation(s) ending in M satisfy the condition of the lemma with grade g or less. By applying the induction hypothesis, we can replace these subderivations with good ones. By asymmetry, M is not now badly misplaced with respect to R, nor can any of the other inferences be badly misplaced with respect to R, since they were not so in the original derivation. It follows that the result is a good derivation. ■

Now, continuing the proof of Theorem 2, define the *reluctance* of D to be the number of rule applications R such that the subderivation D_R of D rooted in R is not good. We proceed by induction on reluctance. If reluctance is zero, D is itself good.

Now suppose the theorem holds for derivations of reluctance d, and consider D of reluctance $d+1$. Since D is finite, there must be a highest inference R such that some inference is badly misplaced with respect to R in the subderivation D_R rooted at R. This D_R satisfies the condition of Lemma 7. Therefore this D_R can be replaced with a corresponding eager derivation, giving a new derivation of smaller reluctance. The induction hypothesis then shows that the resulting derivation can be made eager. ■

B Proof of Theorem 3

Let Γ and Δ be multisets of tracked prefixed expressions in which each formula is tracked by the empty set and prefixed by the empty prefix. There is a proof of $\Gamma \longrightarrow \Delta$ in SCL exactly when there is a proof of $\Gamma; \longrightarrow; \Delta$ in SCLP in which every block is canceled.

Proof. As observed already in Section 2.4, there is an SCL proof of $\Gamma \longrightarrow \Delta$ exactly when there is an SCLI proof of $\Gamma \longrightarrow \Delta$. By Theorem 2 of Section 2.4, there is an SCLI proof of $\Gamma \longrightarrow \Delta$ exactly when there is an *eager* SCLI proof of $\Gamma \longrightarrow \Delta$. By Lemma 3, there is an eager SCLI proof of $\Gamma \longrightarrow \Delta$ exactly when there is an eager articulated SCLI proof of $\Gamma; \longrightarrow; \Delta$. And by Lemma 4, there is an eager articulated SCLI proof of $\Gamma; \longrightarrow; \Delta$ exactly when there is an eager SCLS proof of $\Gamma; \longrightarrow; \Delta$.

Continuing through the argument, By the Contraction Lemma, we may assume without loss of generality that $\Gamma; \longrightarrow; \Delta$ is a simple sequent. We know from its lack of prefixes that the sequent $\Gamma; \longrightarrow; \Delta$ is also spanned and balanced. By Lemma 5 of Section B.2.3, then, there is an eager SCLS proof of $\Gamma; \longrightarrow; \Delta$ exactly when there is a blockwise eager SCLB derivation of $\Gamma; \longrightarrow; \Delta$ in which every block is canceled, linked, isolated, simple, balanced and spanned. And by Lemma 6, there is a blockwise eager SCLB derivation of $\Gamma; \longrightarrow; \Delta$ in which every block is canceled, linked, isolated, simple, balanced and spanned exactly when there is an SCLP derivation of $\Gamma; \longrightarrow; \Delta$ in which every inference is linked. And if every inference is linked, every block is canceled. ■

B.1 Proof of Lemma 4

We show in this section that an articulated SCLI proof with end-sequent $\Pi; \longrightarrow; \Theta$ corresponds to an SCLS proof with end-sequent $\Pi; \longrightarrow; \Theta$, and vice versa. In fact, to transform SCLS to articulated SCLI we have a simple structural induction which replaces $(\supset\to^S)$ with $(\supset\to)$ using the weakening lemma; the soundness of SCLS over SCLI then follows by Lemma 3. Thus, here we are primarily concerned with completeness of a new sequent inference figure.

The use of $(\supset\to^S)$ in eager derivations ensures that the processing of each new goal refers directly to global program statements. To formalize this idea, we introduce the notion of a *fresh* inference.

Definition 33 (Fresh) *Let D be an SCLV derivation. An inference R in D is* fresh *exactly when R is a right inference and the path from R to the root never follows the left spur of any $(\supset\to)$ inference.*

Lemma 8 *Let D be an eager SCLV derivation with an end-sequent of the form*

$$\Pi; \to \Delta; \Theta$$

and consider a subderivation D' of D rooted in a fresh inference R. Then the end-sequent of D' also has the form

$$\Pi'; \to \Delta'; \Theta'$$

for some Π', Δ' and Θ'.

Proof. Suppose otherwise, and consider a maximal D' whose end-sequent contains a non-empty multiset of local statements Γ. We can describe D' equivalently as the subderivation of D that is rooted in a lowest fresh inference R when the end-sequent of D contains some local statements. R cannot be the first inference of D, so there must be an inference S in D immediately below R. If S is a left rule, then the fact that D is eager leads to a contradiction. R must be based in S, or else R will be delayed. This means S is an implication inference; but given that R is fresh, R must appear along the branch of $(\supset\to^S)$ without local statements. Meanwhile, if S is a right rule, it follows from the formulation of the rules that if the end-sequent of D_R has non-empty local statements then the end-sequent of D_L must also. This contradicts the assumption that R is first. ■

Now we proceed with the proof of Lemma 4. We assume an eager SCLV derivation D with such an end-sequent; we show that we can transform it into an eager SCLS derivation D' with the same end-sequent. The proof is by induction on the number of occurrences of $(\supset\to)$ inferences in D.

In the base case, there are no $(\supset\to)$ inferences and D' is just D.

Suppose the claim holds for derivations where $(\supset\to)$ is used fewer than n times, and suppose D is a derivation in which $(\supset\to)$ is used n times. Choose an inference L of $(\supset\to)$ with no other $(\supset\to)$ inference closer to the root of D; we must rewrite the left subderivation at L to match the $(\supset\to^S)$ inference figure. We distinguish a subderivation D' of D as a function of L and draw on the inferences in D' to replace this subderivation—in particular, we identify D' as the largest subderivation of D containing L but no right inferences or segment boundaries below L.

Using Lemma 8, we develop a schema of D' thus:

$$\dfrac{\dfrac{D^A}{\Pi;\Gamma,A\supset B^\mu_X \to A^\mu_X,\Delta;\Theta} \qquad \dfrac{D^B}{\Pi;\Gamma,A\supset B^\mu_X,B^\mu_X \to \Delta;\Theta}}{D^L\left\{\begin{array}{c}\Pi;\Gamma,A\supset B^\mu_X \longrightarrow \Delta;\Theta\\ \vdots\\ \Pi;\longrightarrow \Delta;\Theta\end{array}\right.}\,L$$
$$\overline{\text{(Segment boundary or right rule)}}$$

We suppose L applies to an expression $A\supset B^\mu_X$; the left subderivation of L, D^A adds the goal A; the right, D^B, uses the assumption B. The subderivation of D' from the end-sequent of L abstracts the left inferences performed elsewhere in this segment (and any subgoals that these inferences trigger). We notate this tree of inferences D^L. By Lemma 8, D' ends with a sequent of the form $\Pi;\longrightarrow\Delta;\Theta$. Because of the form of the intervening rules, we have the same succedent $\Delta;\Theta$ at L, as well as the same global statements Π.

We use D^L to construct an eager SCLS derivation $\mathcal{A}$ corresponding to D^A; we will substitute the result for the left subtree at L to revise L to fit the $(\supset\to^S)$ figure. In outline, the derivation we aim for is an eager SCLS version of:

$$\frac{D^A}{D^L + A^\mu_X}$$

The problem is that if D^A is rooted in a right inference to A, we will not obtain an eager derivation when we reassemble L. The SCLS derivation $\mathcal{A}$ we use is actually constructed by recursion on the structure of D^A, applying this kind of transformation at appropriate junctures. At each stage, we call the subderivation of D^A we are considering D'^A.

For the base case, this subderivation is an axiom, and we construct this subderivation as a result. If D'^A ends in a right rule, the construction proceeds inductively by constructing corresponding subderivations and recombining them by the same right rule. With a right inference here, the resulting derivation must be eager since the subderivations are eager.

If D'^A ends in a left inference, the construction is not inductive. We observe that D'^A has an end-sequent of the form

$$\Pi,\Pi';\longrightarrow\Delta,\Delta';\Theta,\Theta'$$

(The inventory of expressions can only be expanded, and that only in certain places, as we follow right inferences to reach D'^A.) So we first weaken D^L by the needed additional expressions—Π' on the left and Δ' (locally) and Θ' (globally) on the right; then we identify the open leaf in D^L with D'^A, obtaining a larger derivation D_I defined as:

$$\frac{D'^A}{\Pi' + D'^L + A^\mu_X + \Delta';\Theta'}$$

Any delayed inference in D_I would in fact be delayed in D'^A, so this is an eager derivation. The result has, moreover, fewer than n $(\supset\to)$ inferences, since it omits at least L from D'. Then the induction hypothesis applies to give the needed SCLS derivation $\mathcal{A}$.

Given the derivation $\mathcal{A}$ so constructed, we substitute $\mathcal{A}$ for D^A in D. The result D^* is an eager derivation; $\mathcal{D}^*$ contains an $(\supset\to^S)$ inference corresponding to L and therefore contains fewer than n uses of $(\supset\to)$. The induction hypothesis applies to transform $\mathcal{D}^*$ to the needed overall derivation. ■

B.2 Proof of Lemma 5

B.2.1 Replacing Herbrand terms

To begin, it is convenient to observe that the use of indexed Herbrand terms allows us to rename Herbrand terms in a proof under certain conditions.

Lemma 9 (Substitution) *Let $\mathcal{D}$ be an SCLU derivation with end-sequent*

$$\Pi; \longrightarrow ; \Theta$$

in which no Herbrand terms or Herbrand prefixes appear; consider a spanned simple subderivation $\mathcal{D}'$ in which a modal Herbrand function η^u_A occurs in some sequent, but does not occur in the end-sequent. Let η^v_A be a Herbrand function that does not occur in $\mathcal{D}$. Then we can construct a proof $\mathcal{D}^*$ containing corresponding inferences in a corresponding order to $\mathcal{D}$ but in which Herbrand terms and Herbrand prefixes are adjusted so that η^v_A is used in place of η^u_A precisely in the subderivation corresponding to $\mathcal{D}'$.

The **proof** is by induction on the structure of derivations. A complex substitution may be required, because the Herbrand calculus may require not only the replacement of η^u_A itself but also the replacement of Herbrand terms that depend indirectly on η^u_A. It is convenient to begin by replacing any first-order Herbrand term not introduced by a $(\exists\to)$ or $(\to\forall)$ inference by a distinguished constant c_0—starting with leaves of the derivation and working downward. This replacement is to ensure that each first-order and modal Herbrand term in $\mathcal{D}$ is determined from an expression in the end-sequent of $\mathcal{D}$ by a finite number of steps of inference. We continue with the systematic replacement of η^u_A and its dependents. In both cases, the form of $\mathcal{D}$ ensures that a finite substitution can systematically rename all these Herbrand terms as required. We use the fact that each sequent is simple and spanned to extend this substitution inductively upward. Because each sequent is spanned the substitution does not need to be extended at $(\Box\to)$ inferences; because each sequent is simple the substitution can be extended freshly at $(\to\Box)$ and $(\to>)$ inferences. Finally, the form of first-order Herbrand terms ensures that a finite extension of the substitution suffices for $(\to\exists)$ and $(\forall\to)$ inferences. ■.

B.2.2 Rectifying blocks

The transformation of individual blocks appeals to the following definition of *required* elements of proofs.

Definition 34 (Required) *Given a derivation $\mathcal{D}$ with end-sequent*

$$\Pi; \Gamma \longrightarrow \Delta; \Theta$$

we say that an expression occurrence E in Θ or Π is required *iff either it is linked or some block in $\mathcal{D}$ is adjacent to the root block and has an end-sequent*

$$\Pi'; \longrightarrow ; \Theta'$$

in which Π' *or* Θ' *contains an expression occurrence based in* E*.*

Lemma 10 (Rectification) *We are given a blockwise eager SCLU derivation* D *such that: every block in* D *is canceled and isolated; every block in* D *other than the root is spanned, linked, balanced and simple; and the end-sequent of* D *is balanced. We transform* D *to an SCLU derivation* D' *in which every block is canceled, linked, isolated, balanced and simple and every block other than the root is spanned. Every block in* D' *other than the root block is identical to a block of* D*; and the inferences in the root block of* D *correspond to inferences in the same order in* D *(and so* D' *is blockwise eager). If the end-sequent of* D *is spanned then* D' *is spanned and isolated.*

Proof. We describe a transformation that establishes the following inductive property given D. There are simple multisets $\Pi_M \subseteq \Pi$ and $\Theta_M \subseteq \Theta$, together with multisets $\Gamma' \subseteq \Gamma$ and $\Delta' \subseteq \Delta$ such that: any Θ' that spans Π_M includes Θ_M; and for any simple Π' with $\Pi_M \subseteq \Pi' \subseteq \Pi$ and any simple Θ' with $\Theta' \subseteq \Theta$ such that Π' and Θ' are spanned by Θ' and the pair Π', Θ' is balanced, there is a D' in which every block is canceled, linked, balanced, balanced and simple, with end-sequent:

$$\Pi'; \Gamma' \longrightarrow \Delta'; \Theta'$$

In this D', each expression in Γ' is linked; each expression in Δ' is linked; each Π_M expression that occurs in Π' is required and each Θ_M expression that occurs in Θ' is linked. Every block in D' other than the root block is identical to a block of D; and the inferences in the root block of D correspond to inferences in the same order in D. Finally, if Γ' and Δ' are spanned by Θ' then D' is spanned; if D is linked then D' contains all the axioms of D.

At axioms, for D of

$$\Pi; \Gamma, A^{\mu}_{X} \longrightarrow A^{\mu}_{Y}, \Delta; \Theta$$

Π_M and Θ_M are empty, while $\Gamma' = A^{\mu}_{X}$ and $\Delta' = A^{\mu}_{X}$. Assume we are given simple Π' from Π and simple Θ' from Θ with Π' and Θ' spanned by Θ'. We construct D' of

$$\Pi'; A^{\mu}_{X} \longrightarrow A^{\mu}_{Y}; \Theta'$$

If A^{μ}_{X} is spanned by Θ', this axiom is spanned too; the remaining conditions are immediate.

At inferences, consider as a representative case $(\vee \rightarrow)$. D ends:

$$\dfrac{\begin{matrix} D_1 \\ \Pi; \Gamma, A \vee B^{\mu}_{X}, A^{\mu}_{X} \longrightarrow \Delta; \Theta \end{matrix} \qquad \begin{matrix} D_2 \\ \Pi; \Gamma, A \vee B^{\mu}_{X}, B^{\mu}_{X} \longrightarrow \Delta; \Theta \end{matrix}}{\Pi; \Gamma, A \vee B^{\mu}_{X} \longrightarrow \Delta; \Theta}$$

The blocks of D_1 and D_2 either contain the root or are blocks from D; the Herbrand prefixes in the end-sequents of D_1 and D_2 occur with the same distribution as in D. Therefore we can apply the induction hypothesis to get Π_{M1}, Θ_{M1}, Γ'_1 and Δ'_1 for D_1; we can apply it to get Π_{M2}, Θ_{M2}, Γ'_2 and Δ'_2 for D_2. To transform D itself, we perform case analysis on Γ'_1 and Γ'_2.

If Γ'_1 does not contain an occurrence of A^{μ}_{X}, then $\Pi_M = \Pi_{M1}$, $\Theta_M = \Theta_{M1}$, $\Gamma' = \Gamma'_1$ and $\Delta' = \Delta'_1$; D'_1 suffices to carry through the induction hypothesis.

Similarly, if Γ'_2 does not contain an occurrence of B^{μ}_{X}, then $\Pi_M = \Pi_{M2}$, $\Theta_M = \Theta_{M2}$, $\Gamma' = \Gamma'_2$ and $\Delta' = \Delta'_2$; D'_2 suffices to carry through the induction hypothesis.

Otherwise, we will set up $\Pi_M = \Pi_{M1} \cup \Pi_{M2}$ and $\Theta_M = \Theta_{M1} \cup \Theta_{M2}$ (as sets); by the inductive characterization of Π_{M1}, Π_{M2}, Θ_{M1} and Θ_{M2}, any Θ' that spans both Π_{M2} and Π_{M2} includes both Θ_{M1} and Θ_{M2}. We also set up Γ' as the multiset containing at least one occurrence of $A \vee B^{\mu}_{X}$ and as many expression occurrences of any expression as either are found in $\Gamma'_1 \backslash A^{\mu}_{X}$ or are found in $\Gamma'_2 \backslash B^{\mu}_{X}$; we set up Δ' as the multiset containing as many expression occurrences of any expression as are found in either Δ'_1 or Δ'_2.

To continue, we now consider simple Π' from Π and simple Θ' from Θ such that $\Pi_{M1} \subseteq \Pi'$, $\Pi_{M2} \subseteq \Pi'$, Π' and Θ' are spanned by Θ', and the pair Π', Θ' is balanced. We know that Θ' includes Θ_M. We can apply the inductive property to transform D_1 and D_2 into derivations with the inductive property:

$$\begin{array}{c} D'_1 \\ \Pi'; \Gamma'_1 \longrightarrow \Delta'_1; \Theta' \end{array} \qquad \begin{array}{c} D'_2 \\ \Pi'; \Gamma'_2 \longrightarrow \Delta'_2; \Theta' \end{array}$$

We weaken *the lowest block* of D'_1 on the left by the expressions in Γ^+ and not already in Γ' and on the right by the expressions in Δ^+ and not already in Δ', giving D^+_1. We similarly weaken the lowest block of D'_2 on the left by the expressions in Γ^+ and not already in Γ'_2 and on the right by the expressions in Δ^+ and not already in Δ'_2, giving D^+_2. Only the lowest blocks are affected by the weakening transformations, so other blocks remain canceled, linked, spanned, isolated and simple; the lowest block in each case remains canceled. The lowest blocks also remain linked since no inferences are added; and they remain simple (and balanced) because no weakening occurs in the global areas. Construct D' as

$$\frac{\begin{array}{c} D^+_1 \\ \Pi'; \Gamma^+, A^{\mu}_{X} \longrightarrow \Delta^+; \Theta' \end{array} \qquad \begin{array}{c} D^+_2 \\ \Pi'; \Gamma^+, B^{\mu}_{X} \longrightarrow \Delta^+; \Theta' \end{array}}{\Pi'; \Gamma^+ \longrightarrow \Delta^+; \Theta'}$$

The end-sequent is simple and balanced so the root block is simple and balanced; the inference is linked since A^{μ}_{X} and B^{μ}_{X} are linked in the subderivations, so the root block is linked. The root block remains canceled as always.

Any Π_M expression is required here because it is required either in D^+_1 in virtue of its membership in Π_{M1} or in D^+_2 in virtue of its membership in Π_{M2}; likewise any Θ_M expression is linked here because it is linked either in D^+_1 in virtue of its membership in Θ_{M1} or in D^+_2 in virtue of its membership in Θ_{M2}. Thus, except for the spanning conditional, we have shown everything we need of this D'.

Finally, then, if Γ' and Δ' is spanned by Θ', Δ'_1 and Δ'_2 are spanned by Θ' and Γ'_1 and Γ'_2 are spanned by Θ' in the resulting (spanned) subderivations D'_1 and D'_2. This shows that the end-sequent of D' is also spanned, so D' itself is spanned.

This reasoning is representative of the construction required also for $(\wedge \rightarrow)$, $(\exists \rightarrow)$, $(\forall \rightarrow)$, $(\rightarrow \wedge)$, $(\rightarrow \vee)$, $(\rightarrow \exists)$, $(\rightarrow \forall)$, (decide) and (restart). It applies also for $(\supset\rightarrow^{S})$, with the obvious caveat that we do not weaken the left subderivation to match local left expressions, since the form of the $(\supset\rightarrow^{S})$ inference requires there to be none.

Next we have $(\vee \rightarrow^B)$; we consider the representative case of $(\vee \rightarrow^B_L)$. D ends:

$$\dfrac{\begin{array}{c}D_1\\ \Pi_0,\Pi;\Gamma,A\vee B^{\mu}_X,A^{\mu}_X \longrightarrow \Delta;\Theta_0,\Theta\end{array} \qquad \begin{array}{c}D_2\\ \Pi_0,B^{\mu}_X;\longrightarrow \Theta_0\end{array}}{\Pi_0,\Pi;\Gamma,A\vee B^{\mu}_X \longrightarrow \Delta;\Theta_0,\Theta}$$

We treat this specially to respect the block boundary before D_2. In particular, we apply the induction hypothesis to D_1 (as we may since its end-sequent has the same distribution of Herbrand prefixes as does that of D), to get Π_{M1}, Θ_{M1}, Γ'_1 and Δ'_1. If A^{μ}_X does not occur in Γ'_1, we let $\Pi_M = \Pi_{M1}$, $\Theta_M = \Theta_{M1}$, $\Gamma' = \Gamma'_1$ and $\Delta' = \Delta'_1$; any derivation D'_1 constructed from appropriate Π' and Θ' suffices to carry through the induction hypothesis.

Otherwise, we get $\Pi_M = \Pi_{M1} \cup \Pi_{e0}$ (as a set), $\Theta_M = \Theta_{M1}$; any Θ' that spans Π_M also spans Π_{M1} and so includes Θ_M. $\Delta' = \Delta'_1$ and Γ' contains Γ'_1 with the occurrence of A^{μ}_X removed, together with an occurrence of $A\vee B^{\mu}_X$ if Γ'_1 does not already contain such an expression.

Assume simple Π' with $\Pi_M \subseteq \Pi' \subseteq \Pi$ and simple Θ' with $\Theta' \subseteq \Theta$ with Π' and Θ' spanned by Θ' and the pair Π',Θ' balanced. As before, we must have Θ_M included in Θ'. We therefore obtain D'_1 by the inductive property; we then weaken D'_1 locally within the lowest block by $A\vee B^{\mu}_X$ on the left if necessary, to obtain a good derivation D^*_1.

The needed D' is now constructed as:

$$\dfrac{\begin{array}{c}D^*_1\\ \Pi';\Gamma',A^{\mu}_X \longrightarrow \Delta';\Theta'\end{array} \qquad \begin{array}{c}D_2\\ \Pi_0,B^{\mu}_X;\longrightarrow \Theta_0\end{array}}{\Pi';\Gamma' \longrightarrow \Delta';\Theta'}$$

We first argue that the construction instantiates the $(\vee \rightarrow^B_L)$ inference rule. Every Herbrand prefix in Π_{0e} and B^{μ}_X occurs in Π' or Γ', so Π_{0e} and B^{μ}_X are spanned by Θ'. But because the root block in D is isolated, Π_{0e} and B^{μ}_X are spanned minimally by Θ_0. Thus $\Theta_0 \subseteq \Theta'$. $\Pi_{0e} \subseteq \Pi_M$ by construction; by isolation Π_0 is the smallest set such that the pair of Π_0,Θ_0 is balanced. But since Π',Θ' is balanced, $\Pi_0 \subseteq \Pi'$.

Now we show that D' so constructed has the needed properties. The end-sequent is simple and balanced so the root block is simple and balanced. The inference is linked: A^{μ}_X is linked in D'_1 by the induction hypothesis; B^{μ}_X is linked in D_2 because D_2 begins a new block which by assumption is canceled. The root block remains canceled as always. Any Π_M expression is required here because either a corresponding expression Π_{0e} in the new block at the left subderivation is based on it, or because it is required in D'_1. Every Θ_M is linked because it is linked in D^*_1.

Finally, if Γ' and Δ' are spanned by Θ', then Δ'_1 and Γ'_1 are spanned by Θ'_1. The new subderivation D'_1 is therefore spanned by the inductive property; this ensures that the overall derivation is spanned.

Next consider $(\Box \rightarrow)$. D ends:

$$\dfrac{\begin{array}{c}D_1\\ \Pi;\Gamma,\Box_i A^{\mu}_X, A^{\mu\nu}_{X,\mu\nu} \longrightarrow \Delta;\Theta\end{array}}{\Pi;\Gamma,\Box_i A^{\mu}_X \longrightarrow \Delta;\Theta}$$

As always, we apply the induction hypothesis to D_1 (as we may since the Herbrand prefixes on Π and Θ formulas remain the same) to obtain Π_{M1}, Θ_{M1}, Γ'_1 and Δ'_1. If $A^{\mu\nu}_{X,\mu\nu}$ does not occur in Γ'_1, we let $\Pi_M = \Pi_{M1}$, $\Theta_M = \Theta_{M1}$, $\Gamma' = \Gamma'_1$ and $\Delta' = \Delta'_1$; any subderivation D'_1 obtained by the inductive property suffices to witness the inductive property for D.

Otherwise we obtain Γ' by extending Γ'_1 by the principal expression $\Box_i A^{\mu}_X$ if necessary and eliminating the side expression $A^{\mu\nu}_{X,\mu\nu}$; $\Pi_M = \Pi_{M1}$, $\Theta_M = \Theta_{M1}$ and $\Delta' = \Delta'_1$. (Since these are common to the subderivation, any Π' that spans Π_M includes Θ_M.) Now we consider Π' with $\Pi_M \subseteq \Pi' \subseteq \Pi$ and Θ' with $\Theta' \subseteq \Theta$, Π' and Θ' spanned by Θ' and the pair Π',Θ' balanced. As always, we have $\Theta_M \subseteq \Theta'$. We obtain D'_1 using Π' and Θ', and weaken the lowest block by local formulas; calling the result D^+_1, we can produce D' by the following construction:

$$\dfrac{\begin{array}{c} D^+_1 \\ \Pi';\Gamma', A^{\mu\nu}_{X,\mu\nu} \longrightarrow \Delta';\Theta' \end{array}}{\Pi';\Gamma' \longrightarrow \Delta';\Theta'}$$

Everything is largely as before. The key new reasoning comes when we assume that Γ' and Δ' are spanned by Θ'. We must argue that $\Gamma', A^{\mu\nu}_{X,\mu\nu}$ is in fact spanned by Θ'. Since $A^{\mu\nu}_{X,\mu\nu}$ is linked in D^+_1, there must be an axiom in this block which is based in $A^{\mu\nu}_{X,\mu\nu}$; indeed, since the expression occurs as a local antecedent, this axiom must occur within the segment. This axiom must pair expressions prefixed by a path μ' where $\mu\nu$ is a prefix of μ'. But because D' remains blockwise eager, no inferences apply to Δ' or Θ' formulas within the segment (nor can they in this fragment augment the Δ' or Θ' formulas within the segment); therefore some Δ' expression is associated with Herbrand prefix μ'. But since Δ' is spanned by Θ', we have that every prefix of μ' is associated with some Θ' expression; so every prefix of $\mu\nu$ is associated with some Θ' expression. Thus D^+_1 is spanned and in turn D' is spanned.

We have one last representative class of inferences in D: $(\rightarrow \Box)$ and $(\rightarrow >)$. We illustrate with the case where D ends in $(\rightarrow >)$:

$$\dfrac{\begin{array}{c} D_1 \\ \Pi, A^{\mu\eta}_{X,\mu\eta};\Gamma \longrightarrow \Delta, A >_i B^{\mu}_X;\Theta, B^{\mu\eta}_{X,\mu\eta} \end{array}}{\Pi;\Gamma \longrightarrow \Delta, A >_i B^{\mu}_X;\Theta}$$

We begin by applying the induction hypothesis to D_1 (as we can, given the symmetric extension of Π and Θ by labeled expressions). We obtain Θ_{M1}, Π_{M1}, Γ'_1 and Δ'_1; we consider alternative cases in response to Θ and Θ_{M1}. First we suppose $B^{\mu\eta}_{X,\mu\eta} \notin \Theta$. It follows by our assumption about D that $A^{\mu\eta}_{X,\mu\eta} \notin \Pi$ either, nor does η occur in Θ. For this case, we start by defining an overall Π_M and Θ_M: Θ_M is Θ_{M1} with any occurrence of $B^{\mu\eta}_{X,\mu\eta}$ eliminated; Π_M is Π_{M1} with any occurrence of $A^{\mu\eta}_{X,\mu\eta}$ eliminated. Π_M contains no occurrences of $\mu\eta$, since Π does not; thus given the inductive property of Θ_{M1} and Π_{M1}, any Θ' that spans Π_M spans Θ_M. We define Γ' and Δ' so that $\Gamma' = \Gamma'_1$ and Δ' contains Δ'_1 together with an occurrence of $A >_i B^{\mu}_X$, provided Δ'_1 does not already contain one and $B^{\mu\eta}_{X,\mu\eta} \in \Theta_{M1}$ or $A^{\mu\eta}_{X,\mu\eta} \in \Pi_{M1}$. So, assume we are given simple Π' with $\Pi_M \subseteq \Pi' \subseteq \Pi$ and simple Θ' with $\Theta' \subseteq \Theta$ (and so $\Theta_M \subseteq \Theta'$) such that Π' and Θ' are spanned by Θ' and the pair Π',Θ' is balanced.

We consider whether $B^{\mu\eta}_{X,\mu\eta} \in \Theta_{M1}$ or $A^{\mu\eta}_{X,\mu\eta} \in \Pi_{M1}$. If neither, we apply the induction hypothesis to D_1 for the case that Θ'_1 is Θ' and Π'_1 is Π'. The resulting derivation D'_1 serves as D'.

Otherwise, $B^{\mu\eta}_{X,\mu\eta} \in \Theta_{M1}$ or $A^{\mu\eta}_{X,\mu\eta} \in \Pi_{M1}$; we apply the inductive property of D_1 for the case that Θ'_1 is $\Theta', B^{\mu\eta}_{X,\mu\eta}$ and Π'_1 is $\Pi', A^{\mu\eta}_{X,\mu\eta}$ (clearly Π'_1 and Θ'_1 are spanned by Θ'_1 assuming Π' and Θ' are spanned by Θ'; the pair Π'_1, Θ'_1 is also balanced given its symmetric extension). If $B^{\mu\eta}_{X,\mu\eta} \in \Theta_{M1}$, by the inductive property it is linked. If $A^{\mu\eta}_{X,\mu\eta} \in \Pi_{M1}$, it is required, but we shall show that it is in fact linked. By the definition of being required, the other possibility is that there is a block adjacent to the root block of D'_1 with end-sequent

$$\Pi'', E; \longrightarrow \Theta''$$

in which the $(\vee \rightarrow^B)$ inference R that bounds the block is based in E and Π'', E or Θ'' contains an expression occurrence based in $A^{\mu\eta}_{X,\mu\eta}$. But since the original block is isolated in the original D, it is E that must be based in $A^{\mu\eta}_{X,\mu\eta}$. But then R is based in $A^{\mu\eta}_{X,\mu\eta}$ and R is linked: in particular its side expression in the left spur) must be linked; so $A^{\mu\eta}_{X,\mu\eta}$ is linked too.

Thus we can weaken D'_1 in its lowest block if necessary by $A >_i B^{\mu}_{X}$ as a local right formula (in Γ), producing D^+_1; D^+_1 remains good by the same argument as the earlier cases. Thus we can construct D' as:

$$\cfrac{\begin{array}{c} D^+_1 \\ \Pi', A^{\mu\eta}_{X,\mu\eta}; \Gamma' \longrightarrow \Delta', A >_i B^{\mu}_{X}; \Theta', B^{\mu\eta}_{X,\mu\eta} \end{array}}{\Pi'; \Gamma' \longrightarrow \Delta'; \Theta'}$$

The end-sequent here is simple and balanced, so the whole root block is simple and balanced. The new inference is linked (in virtue of the linked occurrence of one side expression—$A^{\mu\eta}_{X,\mu\eta}$ or $B^{\mu\eta}_{X,\mu\eta}$) so the whole root block is linked. The root block is of course canceled. Each element of Π_M is required because it is an element of Π_{M1} and required in the immediate subderivation; each element of Θ_M is linked, because it is an element of Θ_{M1} and therefore linked in the immediate subderivation.

To conclude the case, suppose the end-sequent of D is spanned and that Γ' and Δ' are spanned by Θ'; it follows that same property applies to D_1 so the subderivation is spanned. Then the end-sequent must also be spanned.

The alternative case has $B^{\mu\eta}_{X,\mu\eta} \in \Theta$. By assumption, it also has $A^{\mu\eta}_{X,\mu\eta} \in \Pi$. We therefore define an overall Π_M and Θ_M directly as Π_{M1} and Θ_{M1}, respectively; similarly $\Gamma' = \Gamma'_1$ and $\Delta' = \Delta'_1$. To construct the needed D' for appropriate Π' and Θ', we simply apply the induction hypothesis to D_1 for the case that Θ'_1 is Θ' and Π'_1 is Π'. The resulting derivation D'_1 suffices.

Having completed the induction, we argue that we can obtain an overall D' that is isolated, assuming the original D is not only isolated but spanned. Apply the inductive result to D for the case $\Pi' = \Pi$ and $\Theta' = \Theta$; since $\Gamma' \subseteq \Gamma$ and $\Delta' \subseteq \Delta$ we obtain a spanned derivation D' ending

$$\Pi; \Gamma' \longrightarrow \Delta'; \Theta$$

Consider the end-sequent of any block other than the root in D'; it is

$$\Pi_0, E; \longrightarrow; \Theta_0$$

where a corresponding block occurs in D. I argue by contradiction that for any $F \in \Pi_0$ either $F \in \Pi$ or F is based in an occurrence of F as the side expression of an inference in D' in which E is also based. (This will show that D' is isolated.) So consider an exceptional F. Since D is isolated, if $F \notin \Pi$, F is based in an occurrence of F as the side expression of an inference in D in which E is also based; this inference introduces some path symbol η which occurs in the label of F and E. In D', E can not be based in such an inference; otherwise F would also be based in that inference, since D' is simple. (We have assumed that F is not based in such an inference.) But in this case the expression in the end-sequent of D' on which E is based must contain η. Because the end-sequent of D' is spanned the form of Π and Θ is constrained in D, F must occur in Π. This is absurd. ■

We conclude Section B.2.2 by observing some facts about this construction. First, let D' be a derivation obtained by the construction of Lemma 10, and suppose D' is weakened (in a spanned and balanced way) to D'' by adding occurrences of global expressions that either already occur in the end-sequent of D' or never occur as global expressions in D'. Then a straightforward induction shows that D' is obtained again from D'' by the construction of Lemma 10.

Second, observe that if D' is a derivation obtained by the construction of Lemma 10, and D'' is obtained from D'' by the renaming of Herbrand prefixes (as in Lemma 9), then straightforward induction shows that D'' is obtained again from D'' by the construction of Lemma 10.

Third, let D' be a derivation for which the construction of Lemma 10 yields itself. Let ν be a prefix and let the $\Pi;\Theta$ be the smallest balanced pair where Θ contains all the carriers of prefixes of ν introduced in D'. Suppose each expression in Π and Θ has the property that at most one inference of D' has an occurrence of that expression as a side expression. Consider a derivation D'' obtained from D' by weakening globally by Π (on the left) and by Θ (on the right). Let D^* be the result of applying the construction of Lemma 10 to D''. Then D^* contains any subderivation of D' whose end-sequent contains Π and Θ as global formulas. Again this is a straightforward induction; the base case considers a subderivation of D' whose end-sequent contains Π and Θ as global formulas; in this case we apply the first observation. Unary inferences extend the claim immediately. At binary inferences, one subderivation must be unchanged, by the first observation: since Π and Θ are introduced on a unique path, each Π and Θ formula never occurs or already occurs in the end-sequent in that subderivation. Thus the other subderivation necessarily appears in the derivation obtained by the construction of Lemma 10.

B.2.3 Block conversion

We now have the background required to perform the conversion to block structure, and complete the proof of Lemma 5.

We are given a blockwise eager SCLS derivation D whose end-sequent is spanned and balanced and takes the form:

$$\Pi; \longrightarrow; \Theta$$

We can transform D into a blockwise eager SCLB derivation in which every block is canceled, linked, isolated, simple, balanced and spanned.

Proof. Our induction hypothesis is stronger than the lemma. We assume a blockwise eager SCLU derivation D with end-sequent of the form

$$\Pi; \longrightarrow; \Theta$$

in which every block is canceled, linked, isolated, simple, balanced and spanned, such that that the subproof rooted at any $(\vee \rightarrow)$ inference in D is an SCLS derivation. And we identify a distinguished expression occurrence E in the end-sequent of D which is linked. By Lemma 10, it is straightforward to obtain such a derivation from the SCLS derivation (containing only a single block) that we have assumed. We transform D into a blockwise eager SCLB derivation in which every block is canceled, linked, isolated, simple, balanced and spanned and in which E is also linked; we perform induction on the number of $(\vee \rightarrow)$ inferences in D.

In the base case there are no $(\vee \rightarrow)$ inferences, so D itself is an SCLB derivation.

In the inductive case, we assume D with n $(\vee \rightarrow)$ inferences, and assume the hypothesis true for derivations with fewer. We find an application L of $(\vee \rightarrow)$ with no other closer to the root of D. We will transform D to eliminate L.

Let D' denote the smallest subderivation of D containing the full block of D in which L occurs. Explicitly, D' may be D itself; otherwise, D' is rooted at the right subderivation of the highest $(\vee \rightarrow^B)$ inference below L—an inference we will refer to as H. In either case, our assumptions allow us to identify a distinguished linked expression F in the end-sequent of D': either the assumed E from D, or the side expression of the inference H (assumed canceled). Suppose $A \vee B^{\nu}_{Y}$ is the principal of L. We can apply Lemma 9 to rename $A \vee B^{\nu}_{Y}$ to $A \vee B^{\mu}_{X}$ in such a way that each symbol in μ that is introduced in D' is introduced by a unique inference there. Now we can infer the following schema for D':

$$
\begin{array}{c}
\left[\dfrac{\begin{array}{c} D^A \\ \Pi_0, F, \Pi; \Gamma, A \vee B^{\mu}_{X}, A^{\mu}_{X} \longrightarrow \Delta; \Theta_0, \Theta \end{array} \qquad \begin{array}{c} D^B \\ \Pi_0, F, \Pi; \Gamma, A \vee B^{\mu}_{X}, B^{\mu}_{X} \longrightarrow \Delta; \Theta_0, \Theta \end{array}}{\Pi_0, F, \Pi; \Gamma, A \vee B^{\mu}_{X} \longrightarrow \Delta; \Theta_0, \Theta}\, L\right] \\
D^L \\
\Pi_0, F; \longrightarrow ; \Theta_0
\end{array}
$$

That is, the subderivation of D' below L isD^L; the right subderivation above L (in which B is assumed) is D^B; the left is D^A.

We will use the inferences from D^L to construct alternative smaller derivations in place of D^A and D^B. By Θ', indicate the minimal set of formulas required in addition to Θ_0 to span A^{μ}_{X}; by Π' indicate the minimal set of formulas required in addition to Π_0, F and A^{μ}_{X} to ensure that the pair given by $\Pi_0, \Pi', F, A^{\mu}_{X}$ and Θ_0, Θ' is balanced. (This is well-defined because the sequent $\Pi_0, F \longrightarrow \Theta_0$ is already spanned and balanced.) Now we can construct two new subderivations D'^A and D'^B given respectively as follows:

$$
\begin{array}{c}
\left[\dfrac{\begin{array}{c} \Pi' + A^{\mu}_{X} + D^A + \Theta' \\ \Pi_0, F, \Pi, \Pi', A^{\mu}_{X}; \Gamma, A \vee B^{\mu}_{X}, A^{\mu}_{X} \longrightarrow \Delta; \Theta_0, \Theta, \Theta' \end{array}}{\Pi_0, F, \Pi, \Pi', A^{\mu}_{X}; \Gamma, A \vee B^{\mu}_{X} \longrightarrow \Delta; \Theta_0, \Theta, \Theta'}\, \text{decide}\right] \\
\Pi' + A^{\mu}_{X} + D^L + \Theta' \\
\Pi_0, F, \Pi', A^{\mu}_{X}; \longrightarrow ; \Theta_0, \Theta'
\end{array}
$$

$$
\begin{array}{c}
\left[\;\dfrac{\begin{array}{c}[\Pi' + B^\mu_X + D^B + \Theta'] \\ \Pi_0, F, \Pi, \Pi', B^\mu_X; \Gamma, B \vee B^\mu_X, B^\mu_X \longrightarrow \Delta; \Theta_0, \Theta, \Theta'\end{array}}{\Pi_0, F, \Pi, \Pi', B^\mu_X; \Gamma, B \vee B^\mu_X \longrightarrow \Delta; \Theta_0, \Theta, \Theta'}\;\text{decide}\;\right] \\
\Pi' + B^\mu_X + D^L + \Theta' \\
\Pi_0, F, \Pi', B^\mu_X; \longrightarrow ; \Theta_0, \Theta'
\end{array}
$$

That is, we weaken D^A and D^B by global versions of the side expression of inference L throughout their *lowest blocks*; we apply a (decide) inference to obtain a new subderivation to substitute for the subderivation rooted at L in D^L. We weaken by sufficient additional formulas globally in the *lowest blocks* to ensure that the end-sequents of these derivations are balanced and spanned.

Since we have changed only the lowest block here, and have ensured that this block remains isolated and canceled, we can now apply Lemma 10 to obtain corresponding derivations D^A_I and D^B_I in which every block is canceled, linked, isolated, simple, balanced and spanned. In light of our first observation about the construction of Lemma 10, we can see that the inferences of D^A are preserved up to the new (decide) inference. And in light of our third observation about the construction of Lemma 10, given the unique inferences introducing Θ_0 and Π_0, this (decide) inference must be preserved in D^A_I. Thus A^μ_X is linked in D^A_I and for analogous reasons B^μ_X is linked in D^B_I. These derivations satisfy the induction hypothesis as deductions with fewer than n $(\vee \rightarrow)$ inferences; we can apply the induction hypothesis with A^μ_X and B^μ_X as the distinguished linked formulas to preserve. This results in SCLB derivations $\mathcal{A}$ and $\mathcal{B}$ with the same end-sequents as D'^A and D'^B, in which every block is canceled, linked, isolated, simple and spanned, and in which respectively A^μ_X and B^μ_X are linked.

We need only one of $\mathcal{A}$ and $\mathcal{B}$ to reconstruct D' using blocking inferences. For example, we obtain a proof using $(\vee \rightarrow^B_L)$ by using $\mathcal{B}$ in place of D^B as schematized below:

$$
\begin{array}{c}
\left[\;\dfrac{\begin{array}{cc} D^A & \mathcal{B} \\ \Pi_0, F, \Pi; \Gamma, A \vee B^\mu_X, A^\mu_X \longrightarrow \Delta; \Theta_0, \Theta & \Pi_0, F, \Pi', B^\mu_X \longrightarrow \Theta_0, \Theta'\end{array}}{\Pi_0, F, \Pi; \Gamma, A \vee B^\mu_X \longrightarrow \Delta; \Theta_0, \Theta}\;\vee\rightarrow^B_L\right] \\
D^L \\
\Pi_0, F; \longrightarrow ; \Theta_0
\end{array}
$$

In a complementary way, we obtain a proof using $(\vee \rightarrow^B_R)$ by using $\mathcal{A}$ in place of D^A as schematized below:

$$
\begin{array}{c}
\left[\;\dfrac{\begin{array}{cc} D^B & \mathcal{A} \\ \Pi_0, F, \Pi; \Gamma, A \vee B^\mu_X, B^\mu_X \longrightarrow \Delta; \Theta_0, \Theta & \Pi_0, F, \Pi', A^\mu_X \longrightarrow \Theta_0, \Theta'\end{array}}{\Pi_0, F, \Pi; \Gamma, A \vee B^\mu_X \longrightarrow \Delta; \Theta_0, \Theta}\;\vee\rightarrow^B_L\right] \\
D^L \\
\Pi_0, F; \longrightarrow ; \Theta_0
\end{array}
$$

Note that the root block is isolated in both cases, because we have added only as many formulas to Π' and Θ' as are necessary to obtain a balanced, spanned sequent; the remaining expressions originate in the end-sequent of the previous block, which we know was isolated. Thus, in both cases, we have blockwise eager derivations in which every block is canceled, isolated, simple, balanced and spanned, in which fewer than n $(\vee \to)$ inferences are used, and in which only the root block may fail to be linked. We thus need to apply the construction of Lemma 10 again to ensure that the root block is linked. It is possible for the distinguished occurrence of F not to be linked in one of the resulting derivations, but not both. To see this, consider applying the construction of Lemma 10 to $\mathcal{D}'$ itself, as a test: the result will be $\mathcal{D}'$ since $\mathcal{D}'$ is linked. Starting from $\mathcal{D}^A$ and $\mathcal{D}^B$ and axioms elsewhere, each inference in $\mathcal{D}'$ corresponds to an inference in the alternative derivations schematized above. We can argue by straightforward induction that no formula is linked in the reconstructed $\mathcal{D}'$ unless it is also linked in the one of the corresponding reconstructed alternative derivations. And F is linked in $\mathcal{D}'$.

Call the derivation in which F is linked $\mathcal{D}''$; we substitute $\mathcal{D}''$ for $\mathcal{D}'$ in $\mathcal{D}$. Since F remains linked in $\mathcal{D}''$, when we do so, we obtain a blockwise eager SCLU derivation with an appropriate end-sequent, with fewer original $(\vee \to)$ inferences, and in which every block remains canceled, linked, isolated, simple, balanced and spanned, and in which $(\vee \to)$ inferences lie at the root of SCLS derivations. Applying the induction hypothesis to the result gives the required SCLB derivation. ■

B.3 Proof of Lemma 6

We are given a blockwise eager SCLB derivation $\mathcal{D}$, with end-sequent

$$\Pi;\Gamma \longrightarrow \Delta;\Theta$$

in which every block is linked, simple and spanned. We construct an SCLP derivation $\mathcal{D}'$ of which four additional properties hold:

- *the end-sequent of $\mathcal{D}'$ takes the form*

  $$\Pi;\Gamma' \longrightarrow \Delta';\Theta$$

  with $\Gamma' \subseteq \Gamma$ and $\Delta' \subseteq \Delta$;

- *$\mathcal{D}'$ contains in each segment or block all and only the axioms of the corresponding segment or block of $\mathcal{D}$;*

- *whenever $\mathcal{D}'$ contains a sequent of the form*

  $$\Pi^*;\Gamma^* \to F;\Theta^*$$

  F is the only right formula on which an axiom in that block is based; and

- *whenever $\mathcal{D}'$ contains a sequent of the form*

  $$\Pi^*;F \to \Delta^*;\Theta^*$$

  then F is the only left formula on which an axiom in that segment is based.

In the base case, $\mathcal{D}$ is

$$\Pi;\Gamma,A_X^\mu \longrightarrow B_X^\nu,\Delta;\Theta$$

and D' is

$$\Pi; A^{\mu}_{X} \longrightarrow B^{\nu}; \Theta$$

Supposing the claim true for proofs of height h, consider a proof D with height h+1. We consider cases for the different rules with which D could end.

The treatment of $(\rightarrow \wedge)$ is representative of the case analysis for the right rules other than $(\rightarrow >)$. D ends

$$\frac{\Pi; \longrightarrow A^{\mu}_{X}, A \wedge B^{\mu}_{X}, \Delta; \Theta \qquad \Pi; \longrightarrow B^{\mu}_{X}, A \wedge B^{\mu}_{X}, \Delta; \Theta}{\Pi; \longrightarrow A \wedge B^{\mu}_{X}, \Delta; \Theta} \rightarrow \wedge$$

(It is a consequence of Lemma 8 that in the initial derivation there is an empty local area.) We simply apply the induction hypotheses to the immediate subderivations. If the resulting derivations end with (restart), consider the immediate subderivation of the results, otherwise consider the results themselves. These derivations end

$$\Pi; \longrightarrow C; \Theta$$
$$\Pi; \longrightarrow D; \Theta$$

We must have $C = A$; we know from the structure of D that A is linked, and A could not be linked in D unless $C = A$ since D' shows that all of the axioms in D derive from C. For the same reason $D = B$. So we can combine the resulting proofs by an $(\rightarrow \wedge)$ inference to give the needed D'.

The case of $(\rightarrow >)$ proceeds similarly, but relies on an additional observation. D ends

$$\frac{\begin{array}{c} D_1 \\ \Pi, A^{\mu\eta}_{X,\mu\eta}; \longrightarrow \Delta, A >_i B^{\mu}_{X}; B^{\mu\eta}_{X,\mu\eta}, \Theta \end{array}}{\Pi; \longrightarrow \Delta, A >_i B^{\mu}_{X}; \Theta} \rightarrow >$$

We apply the induction hypothesis to D_1 and eliminate any final (restart) inference. This gives us a derivation D_1' of

$$\Pi, A^{\mu\eta}_{X,\mu\eta}; \longrightarrow E; B^{\mu\eta}_{X,\mu\eta}, \Theta$$

If we know that the B-side expression of this inference is linked in this block, then we can conclude, as before, that E is an occurrence of the expression $B^{\mu\eta}_{X,\mu\eta}$. We show this as follows. We know from the structure of D only that *one* of the A-expression and the B-expression must be linked. However, it is straightforward to show that no left expression $A^{\mu\eta}_{X,\mu\eta}$ is linked in an SCLP derivation with a local goal C^{ν}_{Y} unless $\mu\eta$ is a prefix of ν. (The argument is a straightforward variant for example of [Stone, 1999, Lemma 2].) Since D is simple and spanned, η must be new; $B^{\mu\eta}_{X,\mu\eta}$ is the only expression whose associated path term has $\mu\eta$ as a prefix.

Thus, we construct D' using an SCLP inference as

$$\frac{\begin{array}{c} D_1' \\ \Pi, A^{\mu\eta}_{X,\mu\eta}; \longrightarrow B^{\mu\eta}_{X,\mu\eta}; B^{\mu\eta}_{X,\mu\eta}, \Theta \end{array}}{\Pi; \longrightarrow A >_i B^{\mu}_{X}; \Theta} \rightarrow >$$

Now suppose D ends in a left rule other than $(\supset \rightarrow^{S})$ or $(\vee \rightarrow^{B})$. We take $(\wedge \rightarrow)$ as a representative case; then D is:

$$\dfrac{\begin{array}{c} D_1 \\ \Pi;\Gamma,A\wedge B_X^\mu,A_X^\mu,B_X^\mu \longrightarrow \Delta;\Theta \end{array}}{\Pi;\Gamma,A\wedge B_X^\mu \longrightarrow \Delta;\Theta}\wedge\rightarrow$$

Apply the induction hypothesis to D_1. If the result ends in a (decide) inference, let D_1' be the immediate subderivation of the result; otherwise let D_1' be the result itself. D_1' is an SCLP derivation with an end-sequent of the form:

$$\Pi;E \longrightarrow F;\Theta$$

E must be a side expression of the inference in question, here $(\wedge\rightarrow)$; otherwise the corresponding inference could not have been linked in D. One of the inference figures $(\wedge\rightarrow_L)$ and $(\wedge\rightarrow_R)$ must apply depending on which side expression E is. For concrete illustration, we suppose E is A_X^μ; then we construct D' as:

$$\dfrac{\begin{array}{c} D_1' \\ \Pi;A_X^\mu \longrightarrow F;\Theta \end{array}}{\Pi;A\wedge B_X^\mu \longrightarrow F;\Theta}\wedge\rightarrow_L$$

Next, we suppose D ends in $(\supset\rightarrow^S)$, as follows:

$$\dfrac{\begin{array}{c} D_1 \\ \Pi;\longrightarrow A_X^\mu,\Delta;\Theta \end{array} \qquad \begin{array}{c} D_2 \\ \Pi;\Gamma,A\supset B_X^\mu,B_X^\mu \longrightarrow \Delta;\Theta \end{array}}{\Pi;\Gamma,A\supset B_X^\mu \longrightarrow \Delta;\Theta}\supset\rightarrow^S$$

We begin by applying the induction hypothesis to the subderivation D_1. After stripping off any (restart), we obtain an SCLP derivation D_1 with end-sequent

$$\Pi;\longrightarrow C;\Theta$$

By the usual linking argument, the expression C must be identical to A_X^μ. We then apply the induction hypothesis also to the right subderivation. Again, after stripping off any (decide), we get an SCLP derivation D_2 with end-sequent

$$\Pi;D \longrightarrow E;\Theta$$

By the usual linking argument, D must in fact be identical to B_X^μ. Thus we obtain the needed D' by combining the two derivations by the SCLP $(\supset\rightarrow)$ rule:

$$\dfrac{\begin{array}{c} D_1' \\ \Pi;\longrightarrow A_X^\mu;\Theta \end{array} \qquad \begin{array}{c} D_2' \\ \Pi;B_X^\mu \longrightarrow E;\Theta \end{array}}{\Pi;A\supset B_X^\mu \longrightarrow E;\Theta}\supset\rightarrow$$

Finally, for $(\vee\rightarrow^B)$, we consider the representative case of D as schematized below:

$$\dfrac{\begin{array}{c} D_1 \\ \Pi;\Gamma,A_X^\mu \longrightarrow \Delta;\Theta \end{array} \qquad \begin{array}{c} D_2 \\ \Pi',B_X^\mu;\longrightarrow;\Theta' \end{array}}{\Pi;\Gamma,A\vee B_X^\mu \longrightarrow \Delta;\Theta}\vee\rightarrow_L^B$$

We begin by applying the induction hypothesis to D_1, the subderivation in the current block; if necessary, we strip off any initial (decide) inference, obtaining D_1' with an end-sequent that by linking takes the form:

$$\Pi; A^{\mu}_{X} \longrightarrow E; \Theta$$

Next, we apply the induction hypothesis to the other subderivation. Since both local areas are empty in the input subderivation, they remain empty in the result subderivation: this gives D_2' with end-sequent:

$$\Pi', B^{\mu}_{X}; \longrightarrow; \Theta'$$

The two subderivations can be recombined by the SCLP $(\vee \rightarrow_L)$ inference to obtain the needed D':

$$\dfrac{\begin{array}{c} D_1' \\ \Pi; A^{\mu}_{X} \longrightarrow E; \Theta \end{array} \qquad \begin{array}{c} D_2' \\ \Pi', B^{\mu}_{X}; \longrightarrow; \Theta' \end{array}}{\Pi; A \vee B^{\mu}_{X}; \longrightarrow E; \Theta} \vee \rightarrow_L$$

■

References

[Andreoli, 1992] Andreoli, J.-M. (1992). Logic programming with focusing proofs in linear logic. *Journal of Logic and Computation*, 2(3):297–347.

[Auffray and Enjalbert, 1992] Auffray, Y. and Enjalbert, P. (1992). Modal theorem proving: an equational viewpoint. *Journal of Logic and Computation*, 2(3):247–295.

[Baldoni et al., 1993] Baldoni, M., Giordano, L., and Martelli, A. (1993). A multimodal logic to define modules in logic programming. In *ILPS*, pages 473–487.

[Baldoni et al., 1996] Baldoni, M., Giordano, L., and Martelli, A. (1996). A framework for modal logic programming. In Maher, M., editor, *JICSLP 96*, pages 52–66. MIT Press.

[Baldoni et al., 1998a] Baldoni, M., Giordano, L., and Martelli, A. (1998a). A modal extension of logic programming: Modularity, beliefs and hypothetical reasoning. *Journal of Logic and Computation*, 8(5):597–635.

[Baldoni et al., 1998b] Baldoni, M., Giordano, L., and Martelli, A. (1998b). On interaction axioms in multimodal logics: a prefixed tableau calculus. In *Labelled Deduction '98*, Freiburg.

[Basin et al., 1998] Basin, D., Matthews, S., and Viganò, L. (1998). Labelled modal logics: Quantifiers. *Journal of Logic, Language and Information*, 7(3):237–263.

[Beckert and Goré, 1997] Beckert, B. and Goré, R. (1997). Free variable tableaux for propositional modal logics. In *TABLEAUX'97*, LNAI 1227, pages 91–106. Springer.

[Catach, 1991] Catach, L. (1991). TABLEAUX, a general theorem prover for modal logics. *Journal of Automated Reasoning*, 7:489–510.

[Davis, 1994] Davis, E. (1994). Knowledge preconditions for plans. *Journal of Logic and Computation*, 4(5):721–766.

[Debart et al., 1992] Debart, F., Enjalbert, P., and Lescot, M. (1992). Multimodal logic programming using equational and order-sorted logic. *Theoretical Computer Science*, 105:141–166.

[Fariñas del Cerro, 1986] Fariñas del Cerro, L. (1986). MOLOG: A system that extends PROLOG with modal logic. *New Generation Computing*, 4:35–50.

[Fitting, 1972] Fitting, M. (1972). Tableau methods of proof for modal logics. *Notre Dame Journal of Formal Logic*, 13(2).

[Fitting, 1983] Fitting, M. (1983). *Proof Methods for Modal and Intuitionistic Logics*, volume 169 of *Synthese Library*. D. Reidel, Dordrecht.

[Fitting and Mendelsohn, 1998] Fitting, M. and Mendelsohn, R. L. (1998). *First-order Modal Logic*, volume 277 of *Synthese Library*. Kluwer, Dordrecht.

[Frisch and Scherl, 1991] Frisch, A. M. and Scherl, R. B. (1991). A general framework for modal deduction. In *Proceedings of KR*, pages 196–207. Morgan Kaufmann.

[Gabbay, 1985] Gabbay, D. M. (1985). N-Prolog: an extension of Prolog with hypothetical implications, part 2. *Journal of Logic Programming*, 5:251–283.

[Gabbay, 1987] Gabbay, D. M. (1987). Modal and temporal logic programming. In Galton, A., editor, *Temporal Logics and their Applications*, pages 197–237. Academic Press.

[Gabbay, 1992] Gabbay, D. M. (1992). Elements of algorithmic proof. In Abramsky, S., Gabbay, D. M., and Mailbaum, T. S. E., editors, *Handbook of Logic in Theoretical Computer Science*. Oxford.

[Gabbay, 1996] Gabbay, D. M. (1996). *Labelled Deductive Systems*. Oxford.

[Gabbay and Olivetti, 1998] Gabbay, D. M. and Olivetti, N. (1998). Algorithmic proof methods and cut elimintation for implicational logics: Part I: Modal implication. *Studia Logica*, 61:237–280.

[Gabbay and Reyle, 1984] Gabbay, D. M. and Reyle, U. (1984). N-Prolog: an extension of Prolog with hypothetical implications, part 1. *Journal of Logic Programming*, 4:319–355.

[Giordano and Martelli, 1994] Giordano, L. and Martelli, A. (1994). Structuring logic programs: A modal approach. *Journal of Logic Programming*, 21:59–94.

[Girard, 1993] Girard, J.-Y. (1993). On the unity of logic. *Annals of Pure and Applied Logic*, 59:201–217.

[Goré, 1992] Goré, R. (1992). *Cut-free Sequent and Tableau Systems for Propositional Normal Modal Logics*. PhD thesis, University of Cambridge.

[Goré, 1999] Goré, R. (1999). Tableau methods for modal and temporal logics. In D'Agostino, M., Gabbay, D., Hähnle, R., and Posegga, J., editors, *Handbook of Tableau Methods*. Kluwer, Dordrecht.

[Harland, 1994] Harland, J. (1994). A proof-theoretic analysis of goal-directed provability. *Journal of Logic and Computation*, 4(1):69–88.

[Harland, 1997] Harland, J. (1997). On goal-directed provability in classical logic. *Computer Languages*, 23:161–178.

[Harland et al., 2000] Harland, J., Lutovac, T., and Winikoff, M. (2000). Goal-directed proof search in multiple-conclusioned intuitionistc logic. In *Proceedings of the First International Conference on Computational Logic*, volume LNAI 1861, pages 254–268. Springer.

[Hintikka, 1971] Hintikka, J. (1971). Semantics for propositional attitudes. In Linsky, editor, *Reference and Modality*, pages 145–167. Oxford.

[Hodas and Miller, 1994] Hodas, J. S. and Miller, D. (1994). Logic programming in a fragment of intuitionistic linear logic. *Information and Computation*, 110(2):327–365.

[Howard, 1980] Howard, W. A. (1980). The formulae-as-types notion of construction. In *To H. B. Curry: essays on combinatory logic, lambda calculus, and formalism*, pages 479–490. Academic Press, New York.

[Jackson and Reichgelt, 1987] Jackson, P. and Reichgelt, H. (1987). A general proof method for first-order modal logic. In *Proceedings of IJCAI*, pages 942–944.

[Kleene, 1951] Kleene, S. C. (1951). Permutation of inferences in Gentzen's calculi LK and LJ. In *Two papers on the predicate calculus*, pages 1–26. American Mathematical Society, Providence, RI.

[Kobayashi et al., 1999] Kobayashi, N., Shimizu, T., and Yonezawa, A. (1999). Distributed concurrent linear logic programming. *Theoretical Computer Science*, 227:185–220.

[Kripke, 1963] Kripke, S. A. (1963). Semantical analysis of modal logic. I. Normal modal propositional calculi. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 9:67–96.

[Lincoln and Shankar, 1994] Lincoln, P. D. and Shankar, N. (1994). Proof search in first-order linear logic and other cut-free sequent calculi. In *LICS*, pages 282–291.

[Lobo et al., 1992] Lobo, J., Minker, J., and Rajasekar, A. (1992). *Foundations of Disjunctive Logic Programming*. MIT Press.

[Loveland, 1991] Loveland, D. W. (1991). Near-horn Prolog and beyond. *Journal of Automated Reasoning*, 7:1–26.

[Massacci, 1998a] Massacci, F. (1998a). Single step tableaux for modal logics. Technical Report DIS TR-04-98, Unviersity of Rome "La Sapienza".

[Massacci, 1998b] Massacci, F. (1998b). Single step tableaux for modal logics: methodology, computations, algorithms. Technical Report TR-04, DIS, University of Rome "La Sapienza".

[McCarthy, 1993] McCarthy, J. (1993). Notes on formalizing context. In *IJCAI*, pages 555–560.

[McCarthy, 1997] McCarthy, J. (1997). Modality, si! modal logic, no! *Studia Logica*, 59(1):29–32.

[McCarthy and Buvač, 1994] McCarthy, J. and Buvač, S. (1994). Formalizing context (expanded notes). Technical Report STAN-CS-TN-94-13, Stanford University.

[Miller, 1989] Miller, D. (1989). A logical analysis of modules in logic programming. *Journal of Logic Programming*, 6(1–2):79–108.

[Miller, 1994] Miller, D. (1994). A multiple-conclusion meta-logic. In Abramsky, S., editor, *Proceedings of the International Symposium on Logics in Computer Science*, pages 272–281.

[Miller, 1996] Miller, D. (1996). Forum: A multiple-conclusion specification logic. *Theoretical Computer Science*, 165:201–232.

[Miller et al., 1991] Miller, D., Nadathur, G., Pfenning, F., and Scedrov, A. (1991). Uniform proofs as a foundation for logic programming. *Annals of Pure and Applied Logic*, 51:125–157.

[Moore, 1985] Moore, R. C. (1985). A formal theory of knowledge and action. In Hobbs, J. R. and Moore, R. C., editors, *Formal Theories of the Commonsense World*, pages 319–358. Ablex, Norwood NJ.

[Morgenstern, 1987] Morgenstern, L. (1987). Knowledge preconditions for actions and plans. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pages 867–874, Milan Italy.

[Nadathur, 1998] Nadathur, G. (1998). Uniform provability in classical logic. *Journal of Logic and Computation*, 8(2):209–229.

[Nadathur and Loveland, 1995] Nadathur, G. and Loveland, D. W. (1995). Uniform proofs and disjunctive logic programming. In *LICS*, pages 148–155.

[Nonnengart, 1993] Nonnengart, A. (1993). First-order modal logic theorem proving and functional simulation. In *Proceedings of IJCAI*, pages 80–87.

[Ohlbach, 1991] Ohlbach, H. J. (1991). Semantics-based translation methods for modal logics. *Journal of Logic and Computation*, 1(5):691–746.

[Ohlbach, 1993] Ohlbach, H. J. (1993). Optimized translation of multi modal logic into predicate logic. In Voronkov, A., editor, *Logic Programming and Automated Reasoning*, volume 698 of *LNCS*, pages 253–264. Springer, Berlin.

[Orgun and Wadge, 1992] Orgun, M. A. and Wadge, W. W. (1992). Towards a unified theory of intensional logic programming. *Journal of Logic Programming*, 13(4):413–440.

[Otten and Kreitz, 1996] Otten, J. and Kreitz, C. (1996). T-string-unification: unifying prefixes in non-classical proof methods. In *TABLEAUX 96*, volume 1071 of *LNAI*, pages 244–260, Berlin. Springer.

[Prior, 1967] Prior, A. N. (1967). *Past, Present and Future*. Clarendon Press, Oxford.

[Pym and Harland, 1994] Pym, D. and Harland, J. (1994). A uniform proof-theoretic investigation of linear logic programming. *Journal of Logic and Computation*, 4:175–207.

[Robinson, 1965] Robinson, J. A. (1965). A machine oriented logic based on the resolution principle. *Journal of the ACM*, 12(1):23–45.

[Sakakibara, 1987] Sakakibara, Y. (1987). Programming in modal logic: An extension of PROLOG based on modal logic. In Wada, E., editor, *Logic Programming '86*, number 264 in LNCS, pages 81–91. Springer.

[Scherl and Levesque, 1993] Scherl, R. B. and Levesque, H. J. (1993). The frame problem and knowledge-producing actions. In *AAAI*, pages 689–695.

[Schild, 1991] Schild, K. (1991). A correspondence theory for terminological logics: preliminary report. In *IJCAI*, pages 466–471.

[Schmidt, 1998] Schmidt, R. A. (1998). E-Unification for subsystems of S4. In *Rewriting Techniques and Applications*.

[Stone, 1998a] Stone, M. (1998a). Abductive planning with sensing. In *AAAI*, pages 631–636, Madison, WI.

[Stone, 1998b] Stone, M. (1998b). *Modality in Dialogue: Planning, Pragmatics and Computation*. PhD thesis, University of Pennsylvania.

[Stone, 1999] Stone, M. (1999). Representing scope in intuitionistic deductions. *Theoretical Computer Science*, 211(1–2):129–188.

[Stone, 2000] Stone, M. (2000). Towards a computational account of knowledge, action and inference in instructions. *Journal of Language and Computation*, 1:231–246.

[Voronkov, 1996] Voronkov, A. (1996). Proof-search in intuitionistic logic based on constraint satisfaction. In *TABLEAUX 96*, volume 1071 of *LNAI*, pages 312–329, Berlin. Springer.

[Wallen, 1990] Wallen, L. A. (1990). *Automated Proof Search in Non-Classical Logics: Efficient Matrix Proof Methods for Modal and Intuitionistic Logics*. MIT Press, Cambridge.