

# Bidirectional Decision Procedures for the Intuitionistic Propositional Modal Logic **IS4**

Samuli Heilala and Brigitte Pientka

School of Computer Science, McGill University, Montreal, Canada  
{sheilal,bpientka}@cs.mcgill.ca

**Abstract.** We present a multi-context focused sequent calculus whose derivations are in bijective correspondence with normal natural deductions in the propositional fragment of the intuitionistic modal logic **IS4**. This calculus, suitable for the enumeration of normal proofs, is the starting point for the development of a sequent calculus-based bidirectional decision procedure for propositional **IS4**. In this system, relevant derived inference rules are constructed in a forward direction prior to proof search, while derivations constructed using these derived rules are searched over in a backward direction. We also present a variant which searches directly over normal natural deductions. Experimental results show that on most problems, the bidirectional prover is competitive with both conventional backward provers using loop-detection and inverse method provers, significantly outperforming them in a number of cases.

## 1 Introduction

Intuitionistic modal logics are constructive logics incorporating operators of necessity ( $\Box$ ) and possibility ( $\Diamond$ ). Fitch [7], Prawitz [16], Satre [18], and more recently Simpson [19], Bierman and de Paiva [1], and Pfenning and Davies [15] have investigated a broad range of proof-theoretical properties of various logics of this kind. Recently, such logics have also found applications in hardware verification [6] and proposed type systems for staged computation [3] and distributed computing [13]. A logic frequently used in these settings is either the intuitionistic variant of the classical modal logic **S4**, which we will call **IS4**, or a logic that can be expressed through **IS4**, such as Fairtlough and Mendler’s lax logic [6] (see for instance [15] for the relationship between **IS4** and lax logic).

In this light, it is surprising that proof search in **IS4** has not received more attention. Howe has investigated proof enumeration and theorem proving in lax logic [12] and, coming closer to our work, has presented a backward decision procedure for the fragment of propositional **IS4** without the possibility modality [11]. His system performs loop-detection using a history mechanism, but is encumbered by a large number of rules and related provisos (21 axioms and inference rules). It would only grow with the addition of the possibility modality, which would also require a different loop-detection mechanism.

Our contributions begin with a sequent calculus for propositional **IS4** suitable for the enumeration of normal proofs. This forms the basis for the development of a sequent calculus-based *bidirectional IS4* decision procedure, in which derived inference rules relevant to the query are constructed in a forward direction prior to proof search, while derivations constructed using these derived rules are searched over in a backward direction. We also demonstrate that this approach corresponds very closely to an elegant bidirectional decision procedure that searches directly over normal natural deductions. The key to our theoretical justification of both of these decision procedures is a refinement of the well-known subformula property, which we use to restrict nondeterminism in focused proof search in the presence of multiple contexts. To evaluate our approach empirically, we have put together a set of 50 benchmark formulas for **IS4**. Experimental results show that on most problems, the bidirectional prover is competitive with both conventional backward provers using loop-detection and inverse method provers, significantly outperforming them in a number of cases. Although we concentrate on propositional **IS4** in this paper, we believe that the techniques presented are general enough to find applications in other constructive logics, such as the contextual modal logic of Nanevski, Pfenning, and Pientka [14]. Finally, while this paper contains only proof sketches of many of our results, we provide the full proofs in the accompanying technical report [10].

In Sect. 2 we summarize the relevant background and introduce our core natural deduction formalism, while Sect. 3 presents corresponding sequent calculi for proof search in both backward and forward directions, followed by a more detailed discussion of some of the intricacies of focused forward proof search. In Sects. 4 and 5 we describe our bidirectional decision procedure in both a sequent calculus and a natural deduction setting. Experimental results are given in Sect. 6, while Sect. 7 concludes with related and future work.

## 2 Natural Deduction

Formulas in the propositional fragment of **IS4** are given by the grammar

$$A ::= P \mid \perp \mid A \supset A \mid A \wedge A \mid A \vee A \mid \Box A \mid \Diamond A$$

where  $P$  is taken from a countable set of atomic propositional constants and negation and truth are defined notationally in the usual way. Our starting point is a multi-context natural deduction formulation for **IS4** similar to ones proposed by Pfenning and Davies [15] and Bierman and de Paiva [1], except that we impose a restriction that only natural deductions in normal form can be constructed. This is achieved by annotating judgements with their intended direction of reasoning:

$$\begin{array}{ll} \Delta; \Gamma \vdash A \uparrow & A \text{ has a normal proof under hypotheses } \Delta \text{ and } \Gamma, \\ \Delta; \Gamma \vdash A \downarrow & A \text{ can be extracted from hypotheses in } \Delta \text{ and } \Gamma \text{ using} \\ & \text{only elimination rules,} \end{array}$$

$$\begin{array}{c}
\frac{}{\Delta; \Gamma_1, A, \Gamma_2 \vdash A \downarrow} \text{hyp}_1 \quad \frac{}{\Delta_1, A, \Delta_2; \Gamma \vdash A \downarrow} \text{hyp}_2 \quad \frac{\Delta; \Gamma \vdash \perp \downarrow}{\Delta; \Gamma \vdash C \uparrow} \perp\text{E} \\
\frac{\Delta; \Gamma, A_1 \vdash A_2 \uparrow}{\Delta; \Gamma \vdash A_1 \supset A_2 \uparrow} \supset\text{I} \quad \frac{\Delta; \Gamma \vdash A_1 \supset A_2 \downarrow \quad \Delta; \Gamma \vdash A_1 \uparrow}{\Delta; \Gamma \vdash A_2 \downarrow} \supset\text{E} \\
\frac{\Delta; \Gamma \vdash A_1 \uparrow \quad \Delta; \Gamma \vdash A_2 \uparrow}{\Delta; \Gamma \vdash A_1 \wedge A_2 \uparrow} \wedge\text{I} \quad \frac{\Delta; \Gamma \vdash A_1 \wedge A_2 \downarrow}{\Delta; \Gamma \vdash A_j \downarrow} \wedge\text{E}_j \\
\frac{\Delta; \Gamma \vdash A_j \uparrow}{\Delta; \Gamma \vdash A_1 \vee A_2 \uparrow} \vee\text{I}_j \quad \frac{\Delta; \Gamma \vdash A_1 \vee A_2 \downarrow \quad \Delta; \Gamma, A_1 \vdash C \uparrow \quad \Delta; \Gamma, A_2 \vdash C \uparrow}{\Delta; \Gamma \vdash C \uparrow} \vee\text{E} \\
\frac{\Delta; \cdot \vdash A \uparrow}{\Delta; \Gamma \vdash \Box A \uparrow} \Box\text{I} \quad \frac{\Delta; \Gamma \vdash \Box A \downarrow \quad \Delta, A; \Gamma \vdash C \uparrow}{\Delta; \Gamma \vdash C \uparrow} \Box\text{E} \\
\frac{\Delta; \Gamma \vdash A \uparrow}{\Delta; \Gamma \vdash \Diamond A \uparrow} \Diamond\text{I} \quad \frac{\Delta; \Gamma \vdash \Diamond A \downarrow \quad \Delta; A \vdash \Diamond C \uparrow}{\Delta; \Gamma \vdash \Diamond C \uparrow} \Diamond\text{E} \quad \frac{\Delta; \Gamma \vdash A \downarrow \quad A \text{ is atomic}}{\Delta; \Gamma \vdash A \uparrow} \uparrow\downarrow \\
j \in \{1, 2\}
\end{array}$$

**Fig. 1.**  $\mathbf{NJ}_{\mathbf{IS4}}^{\mathbf{N}}$

where  $\Gamma = A_1, \dots, A_n$  is a context of true hypotheses and  $\Delta = B_1, \dots, B_m$  is a modal context of *valid* hypotheses. Valid hypotheses are hypotheses whose truth does not depend on the truth of other formulas, that is, hypotheses that are in some sense “always” or necessarily true. The resulting system, which we will call  $\mathbf{NJ}_{\mathbf{IS4}}^{\mathbf{N}}$ , is shown in Fig. 1. Although the contexts of this system are formally ordered lists, we can afford to be flexible with them, as  $\mathbf{NJ}_{\mathbf{IS4}}^{\mathbf{N}}$  has the usual structural properties of weakening, contraction, and exchange for both contexts. For convenience, we will generally think of contexts in  $\mathbf{NJ}_{\mathbf{IS4}}^{\mathbf{N}}$  as multisets.

The inference rules of  $\mathbf{NJ}_{\mathbf{IS4}}^{\mathbf{N}}$  are largely standard, but to glean some intuition about the modal rules and the two contexts, it is useful to think of the modalities as quantifying truth over worlds in some universe, with some reachability relation defined on the worlds. To say that  $\Box A$  is true is to say that  $A$  is true in *all* worlds reachable from the current one, while to say that  $\Diamond A$  is true is to say that  $A$  is true in *some* world reachable from the current one. The current world represents the environment in which the provability of the succedent is to be established. Under this interpretation, the hypotheses in the modal context can be used in all reachable worlds, while those in the regular context can only be used in the current world.

Note that while  $\mathbf{NJ}_{\mathbf{IS4}}^{\mathbf{N}}$  defines the normal forms that we are interested in during proof search, an unrestricted variant  $\mathbf{NJ}_{\mathbf{IS4}}$  can be obtained by dropping the arrow annotations and the rule  $\uparrow\downarrow$ . In the accompanying technical report [10], we show that the two systems  $\mathbf{NJ}_{\mathbf{IS4}}$  and its “normalized” cousin  $\mathbf{NJ}_{\mathbf{IS4}}^{\mathbf{N}}$  are equivalent in terms of provability. For the interested reader, we also provide a common Hilbert-style axiomatization of  $\mathbf{IS4}$  in [10], along with a proof that the unrestricted system  $\mathbf{NJ}_{\mathbf{IS4}}$  and the axiomatization are equivalent. This lends

support to the claim that we are indeed dealing with the intuitionistic variant of **S4**. Finally, we would like to point out that the separation between modal and ordinary hypotheses is not strictly necessary. Building on work by Bierman and de Paiva [1], we obtain a faithful embedding into a single-context system simply by providing every valid hypothesis with a  $\Box$  operator and merging the two contexts. The details of this embedding are beyond the scope of this paper, but are presented in [10].

While Girard, Lafont, and Taylor suggest that we should think of natural deductions as the “true ‘proof’ objects” [9], natural deduction systems have traditionally not seen much use as formalisms for proof search, mainly as a result of their lack of syntax-directedness. Although we will ultimately return to natural deduction in our search for bidirectional decision procedures, the relationship between backward and forward proof search is perhaps most vividly demonstrated in a sequent calculus setting, which we turn to next.

### 3 Sequent Calculi

Following the approach of Dyckhoff and Pinto [5], we can construct a *focused* sequent calculus for propositional **IS4** whose derivations are in bijective correspondence with normal natural deductions. This system, which we will call **MJ<sub>IS4</sub>**, is shown in Fig. 2 and involves two forms of sequents:

$$\begin{array}{ll} \Delta; \Gamma \rightarrow C & C \text{ can be proved from assumptions } \Delta, \Gamma, \\ \Delta; \Gamma \triangleright A \rightarrow C & C \text{ can be proved from assumptions } \Delta, \Gamma, A, \text{ focusing on} \\ & \text{the assumption } A. \end{array}$$

If a sequent is focused on a formula  $A$ , then the only applicable rules are those with  $A$  as a principal formula. Following Girard [8], we will call the position of the focused formula the *stoup*. As in the natural deduction formulations, contexts in **MJ<sub>IS4</sub>** are technically ordered lists, but the usual structural properties of weakening, exchange, and contraction hold here as well, so an interpretation of contexts as multisets is reasonable. The following key result establishes the close correspondence between **MJ<sub>IS4</sub>** and **NJ<sub>IS4</sub><sup>N</sup>**. The soundness and completeness of **MJ<sub>IS4</sub>** with respect to **NJ<sub>IS4</sub><sup>N</sup>** follow from it.

**Theorem 1 (Bijection between **MJ<sub>IS4</sub>** and **NJ<sub>IS4</sub><sup>N</sup>** derivations).** *Derivations of unfocused sequents in **MJ<sub>IS4</sub>** correspond bijectively to derivations of  $\uparrow$  judgements in **NJ<sub>IS4</sub><sup>N</sup>**.*

*Proof.* We define functions mapping derivations from **NJ<sub>IS4</sub><sup>N</sup>** to **MJ<sub>IS4</sub>** and vice versa. Inductive arguments on the structures of the argument derivations show that the functions are bijections.  $\square$

Although **MJ<sub>IS4</sub>** is suitable for proof search in a backward direction, a naive approach still requires loop-detection to achieve a decision procedure. We will not pursue this direction further here, but instead concentrate on forward proof

$$\begin{array}{c}
\frac{A \text{ is atomic}}{\Delta; \Gamma \triangleright A \rightarrow A} \text{ init} \quad \frac{}{\Delta; \Gamma \triangleright \perp \rightarrow C} \perp\text{L} \\
\frac{\Delta; \Gamma_1, A, \Gamma_2 \triangleright A \rightarrow C}{\Delta; \Gamma_1, A, \Gamma_2 \rightarrow C} \text{ ch}_1 \quad \frac{\Delta_1, A, \Delta_2; \Gamma \triangleright A \rightarrow C}{\Delta_1, A, \Delta_2; \Gamma \rightarrow C} \text{ ch}_2 \\
\frac{\Delta; \Gamma, A_1 \rightarrow A_2}{\Delta; \Gamma \rightarrow A_1 \supset A_2} \supset\text{R} \quad \frac{\Delta; \Gamma \rightarrow A_1 \quad \Delta; \Gamma \triangleright A_2 \rightarrow C}{\Delta; \Gamma \triangleright A_1 \supset A_2 \rightarrow C} \supset\text{L} \\
\frac{\Delta; \Gamma \rightarrow A_1 \quad \Delta; \Gamma \rightarrow A_2}{\Delta; \Gamma \rightarrow A_1 \wedge A_2} \wedge\text{R} \quad \frac{\Delta; \Gamma \triangleright A_j \rightarrow C}{\Delta; \Gamma \triangleright A_1 \wedge A_2 \rightarrow C} \wedge\text{L}_j \\
\frac{\Delta; \Gamma \rightarrow A_j}{\Delta; \Gamma \rightarrow A_1 \vee A_2} \vee\text{R}_j \quad \frac{\Delta; \Gamma, A_1 \rightarrow C \quad \Delta; \Gamma, A_2 \rightarrow C}{\Delta; \Gamma \triangleright A_1 \vee A_2 \rightarrow C} \vee\text{L} \\
\frac{\Delta; \cdot \rightarrow A}{\Delta; \Gamma \rightarrow \Box A} \Box\text{R} \quad \frac{\Delta, A; \Gamma \rightarrow C}{\Delta; \Gamma \triangleright \Box A \rightarrow C} \Box\text{L} \\
\frac{\Delta; \Gamma \rightarrow A}{\Delta; \Gamma \rightarrow \Diamond A} \Diamond\text{R} \quad \frac{\Delta; A \rightarrow \Diamond C}{\Delta; \Gamma \triangleright \Diamond A \rightarrow \Diamond C} \Diamond\text{L} \\
j \in \{1, 2\}
\end{array}$$

**Fig. 2.**  $\mathbf{MJ}_{\text{IS4}}$

search, and on how we can combine ideas from backward and forward search to perform bidirectional proof search.

Constructing  $\mathbf{MJ}_{\text{IS4}}$  proofs in a forward direction — from the top down — is complicated by the presence of multiple contexts, making  $\mathbf{MJ}_{\text{IS4}}$  less than ideal for forward proof search. All  $\mathbf{MJ}_{\text{IS4}}$  derivations begin, at the leaves, with focused sequents of the form  $\Delta; \Gamma \triangleright A \rightarrow A$ , with  $A$  atomic. After a sequence of (possibly zero) left-rule applications, the stoup formula is dropped from the stoup into one of the contexts by an application of  $\text{ch}_1$  or  $\text{ch}_2$ . In a focused forward calculus used as the basis for the inverse method [4], we would proceed in a similar way, but it is not clear which context a stoup formula should be dropped into.

To address this uncertainty, we refine the idea of focusing and develop the system  $\mathbf{MJ}_{\text{IS4}}^{\text{F}}$ , which is suitable for forward proof search and features sequents of three kinds, involving both modal and nonmodal stoups:

$$\begin{array}{ll}
\Delta; \Gamma \mapsto C & C \text{ can be proved using all assumptions in } \Delta, \Gamma, \\
\Delta; \Gamma \triangleright A \mapsto C & C \text{ can be proved using all assumptions in } \Delta, \Gamma, A, \text{ with} \\
& A \text{ assumed true,} \\
\Delta; \Gamma \triangleright \triangleright A \mapsto C & C \text{ can be proved using all assumptions in } \Delta, \Gamma, A, \text{ with} \\
& A \text{ assumed valid.}
\end{array}$$

Note that the forms of the focused sequents reveal which context the stoup formula will drop into. For brevity, we write  $\Delta; \Gamma \triangleright^i A \mapsto C$ ,  $i \in \{1, 2\}$  for either form of focused sequent.

$$\begin{array}{c}
\frac{A \text{ is atomic}}{;\cdot \triangleright^i A \mapsto A} \text{init}_i \quad \frac{}{;\cdot \triangleright^i \perp \mapsto C} \perp\text{L}_i \quad \frac{\Delta; \Gamma \triangleright A \mapsto C}{\Delta; \Gamma, A \mapsto C} \text{ch}_1 \quad \frac{\Delta; \Gamma \triangleright \triangleright A \mapsto C}{\Delta, A; \Gamma \mapsto C} \text{ch}_2 \\
\frac{\Delta; \Gamma, A_1 \mapsto A_2}{\Delta; \Gamma \mapsto A_1 \supset A_2} \supset\text{R}_1 \quad \frac{\Delta; \Gamma \mapsto A_2}{\Delta; \Gamma \mapsto A_1 \supset A_2} \supset\text{R}_2 \\
\frac{\Delta_1; \Gamma_1 \mapsto A_1 \quad \Delta_2; \Gamma_2 \triangleright^i A_2 \mapsto C}{\Delta_1, \Delta_2; \Gamma_1, \Gamma_2 \triangleright^i A_1 \supset A_2 \mapsto C} \supset\text{L}_i \\
\frac{\Delta_1; \Gamma_1 \mapsto A_1 \quad \Delta_2; \Gamma_2 \mapsto A_2}{\Delta_1, \Delta_2; \Gamma_1, \Gamma_2 \mapsto A_1 \wedge A_2} \wedge\text{R} \quad \frac{\Delta; \Gamma \triangleright^i A_j \mapsto C}{\Delta; \Gamma \triangleright^i A_1 \wedge A_2 \mapsto C} \wedge\text{L}_{i,j} \\
\frac{\Delta; \Gamma \mapsto A_j}{\Delta; \Gamma \mapsto A_1 \vee A_2} \vee\text{R}_j \quad \frac{\Delta_1; \Gamma_1, A_1 \mapsto C \quad \Delta_2; \Gamma_2, A_2 \mapsto C}{\Delta_1, \Delta_2; \Gamma_1, \Gamma_2 \triangleright^i A_1 \vee A_2 \mapsto C} \vee\text{L}_i \\
\frac{\Delta; \cdot \mapsto A}{\Delta; \cdot \mapsto \Box A} \Box\text{R} \quad \frac{\Delta, A; \Gamma \mapsto C}{\Delta; \Gamma \triangleright^i \Box A \mapsto C} \Box\text{L}_i \quad \frac{\Delta; \Gamma \mapsto A}{\Delta; \Gamma \mapsto \Diamond A} \Diamond\text{R} \quad \frac{\Delta; A \mapsto \Diamond C}{\Delta; \cdot \triangleright^i \Diamond A \mapsto \Diamond C} \Diamond\text{L}_i \\
i, j \in \{1, 2\}
\end{array}$$

**Fig. 3.**  $\mathbf{MJ}_{\text{IS4}}^{\text{F}}$

The inference rules of  $\mathbf{MJ}_{\text{IS4}}^{\text{F}}$ , shown in Fig. 3, are obtained by reinterpreting the rules of  $\mathbf{MJ}_{\text{IS4}}$  in a forward fashion and by defining the  $\text{ch}_i$  rules to behave as sketched above. The contexts of  $\mathbf{MJ}_{\text{IS4}}^{\text{F}}$ , however, are interpreted differently, in that sequents  $\Delta; \Gamma \mapsto C$  and  $\Delta; \Gamma \triangleright^i A \mapsto C$ ,  $i \in \{1, 2\}$  assert that *all* assumptions in  $\Delta$  and  $\Gamma$ , as well as  $A$  if the sequent is focused, are needed to prove  $C$ . General weakening, which holds in  $\mathbf{MJ}_{\text{IS4}}$ , is thus disallowed, but local weakening is incorporated in the rule  $\supset\text{R}_2$ . Contexts in  $\mathbf{MJ}_{\text{IS4}}^{\text{F}}$  are treated as sets rather than multisets, and we write  $\Gamma_1, \Gamma_2$  and  $\Gamma, A$  for  $\Gamma_1 \cup \Gamma_2$  and  $\Gamma \cup \{A\}$ , respectively.

**Theorem 2 (Soundness and completeness of  $\mathbf{MJ}_{\text{IS4}}^{\text{F}}$  with respect to  $\mathbf{MJ}_{\text{IS4}}$ ).**

1. (Soundness)
  - (a) If  $\Delta; \Gamma \mapsto C$ , then  $\Delta; \Gamma \rightarrow C$ .
  - (b) If  $\Delta; \Gamma \triangleright^i A \mapsto C$ ,  $i \in \{1, 2\}$ , then  $\Delta; \Gamma \triangleright A \rightarrow C$ .
2. (Completeness)
  - (a) If  $\Delta; \Gamma \rightarrow C$ , then  $\Delta'; \Gamma' \mapsto C$  for some  $\Delta' \subseteq \Delta$ ,  $\Gamma' \subseteq \Gamma$ .
  - (b) If  $\Delta; \Gamma \triangleright A \rightarrow C$  and  $A$  is a subformula of a formula in  $\Gamma$ , then either  $\Delta'; \Gamma' \mapsto C$  or  $\Delta'; \Gamma' \triangleright A \mapsto C$  for some  $\Delta' \subseteq \Delta$ ,  $\Gamma' \subseteq \Gamma$ .
  - (c) If  $\Delta; \Gamma \triangleright A \rightarrow C$  and  $A$  is a subformula of a formula in  $\Delta$ , then either  $\Delta'; \Gamma' \mapsto C$  or  $\Delta'; \Gamma' \triangleright \triangleright A \mapsto C$  for some  $\Delta' \subseteq \Delta$ ,  $\Gamma' \subseteq \Gamma$ .

*Proof.* In both cases by simultaneous induction on the structure of the given derivation, using weakening in  $\mathbf{MJ}_{\text{IS4}}$  where necessary.  $\square$

Note that the more fine-grained focusing mechanism of  $\mathbf{MJ}_{\mathbf{IS4}}^{\mathbf{F}}$  could just as well have been introduced in a sequent calculus suitable for backward reasoning, such as  $\mathbf{MJ}_{\mathbf{IS4}}$ . Indeed, the single type of focused sequent in  $\mathbf{MJ}_{\mathbf{IS4}}$  has the role of both types of focused sequents in  $\mathbf{MJ}_{\mathbf{IS4}}^{\mathbf{F}}$ , making the focusing mechanism of  $\mathbf{MJ}_{\mathbf{IS4}}$  in some sense “overloaded”.

The forward calculus  $\mathbf{MJ}_{\mathbf{IS4}}^{\mathbf{F}}$  suggests itself immediately as a basis for an implementation of the inverse method [4], fundamental to which is the classification of the subformulas of a query formula into positive and negative classes. The sign of a subformula determines where in a sequent it may occur (for instance as a goal formula or in the context) and restricts nondeterminism during proof search. We will refine this notion by classifying subformulas as either

1. positive (+) subformulas, which may occur as goal formulas,
2. negative (−) subformulas, which may occur in the nonmodal context,
3. negative focused ( $\sim$ ) subformulas, which may occur in the nonmodal stoup,
4. valid (=) subformulas, which may occur in the modal context, or
5. valid focused ( $\approx$ ) subformulas, which may occur in the modal stoup.

With this intended interpretation, it is straightforward to read the formal definition of refined signed subformulas directly from the inference rules of  $\mathbf{MJ}_{\mathbf{IS4}}^{\mathbf{F}}$ .

**Definition 1 (Signed subformulas).** *A signed subformula  $A^*$  is a formula  $A$  with a sign  $*$   $\in \{+, -, \sim, =, \approx\}$ . The subformula relation  $\leq$  is the smallest reflexive and transitive relation between signed subformulas satisfying the following.*

$$\begin{aligned}
A_1^-, A_2^+ &\leq (A_1 \supset A_2)^+ & A_i^+ &\leq (A_1 \wedge A_2)^+ & A_i^+ &\leq (A_1 \vee A_2)^+ \\
A^+ &\leq (\Box A)^+ & A^+ &\leq (\Diamond A)^+ & A^\sim &\leq A^- \\
A_1^+, A_2^\sim &\leq (A_1 \supset A_2)^\sim & A_i^\sim &\leq (A_1 \wedge A_2)^\sim & A_i^- &\leq (A_1 \vee A_2)^\sim \\
A^\sim &\leq (\Box A)^\sim & A^- &\leq (\Diamond A)^\sim & A^\approx &\leq A^= \\
A_1^+, A_2^\approx &\leq (A_1 \supset A_2)^\approx & A_i^\approx &\leq (A_1 \wedge A_2)^\approx & A_i^- &\leq (A_1 \vee A_2)^\approx \\
A^\approx &\leq (\Box A)^\approx & A^- &\leq (\Diamond A)^\approx \\
&& i &\in \{1, 2\}
\end{aligned}$$

Note that for every negative subformula  $A^-$  of a signed formula  $C^*$ ,  $C^*$  also has, as a subformula, the corresponding negative focused subformula  $A^\sim$ . The converse, however, is not true in general. A similar relation holds for valid and valid focused subformulas. Also, the usual signed subformula property extends to encompass our refined signing scheme, where we write  $\Gamma^-$  and  $\Delta^\approx$  for contexts of signed subformulas of the forms  $A_1^-, \dots, A_n^-$  and  $B_1^\approx, \dots, B_m^\approx$ , respectively.

**Theorem 3 (Signed subformula property).** *Every sequent in an  $\mathbf{MJ}_{\mathbf{IS4}}^{\mathbf{F}}$  derivation of*

$$\Delta^\approx; \Gamma^- \mapsto C^+ \quad \text{or} \quad \Delta^\approx; \Gamma^- \triangleright^i A^* \mapsto C^+, \quad i \in \{1, 2\}$$

where  $*$  is  $\sim$  or  $\approx$  if  $i = 1$  or  $i = 2$ , respectively, is of the form

1.  $D_1^-, \dots, D_n^-; E_1^-, \dots, E_m^- \mapsto F^+$ ,
2.  $D_1^-, \dots, D_n^-; E_1^-, \dots, E_m^- \triangleright E^\sim \mapsto F^+$ , or
3.  $D_1^-, \dots, D_n^-; E_1^-, \dots, E_m^- \triangleright \triangleright D^\approx \mapsto F^+$ ,

where all  $D_j^-$ ,  $E_k^-$ , and  $E^\sim$ ,  $D^\approx$ , and  $F^+$  are signed subformulas of  $\Delta^=$ ,  $\Gamma^-$ ,  $C^+$ , and  $A^*$ .

*Proof.* By simultaneous induction on the structure of the given derivation.  $\square$

Theorem 3 guarantees, for instance, that in any  $\mathbf{MJ}_{\mathbf{IS4}}^F$  derivation of the sequent  $\Delta^=; \Gamma^- \mapsto C^+$ , all leaves are of the forms

$$\frac{A \text{ is atomic}}{;\cdot \triangleright^i A^* \mapsto A^+} \text{init}_i \quad \text{or} \quad \frac{}{;\cdot \triangleright^i \perp^* \mapsto B^+} \perp L_i \quad i \in \{1, 2\}$$

where  $*$  is  $\sim$  or  $\approx$  if  $i = 1$  or  $i = 2$ , respectively, and  $A^*$ ,  $A^+$ ,  $\perp^*$ , and  $B^+$  must be signed subformulas of  $\Delta^=$ ,  $\Gamma^-$ , and  $C^+$ . In general, every rule application considered by an implementation of the inverse method must abide by the conditions set forth by the extended signed subformula property. This provides a foundation for a focused inverse method prover for  $\mathbf{IS4}$  with nondeterminism restricted more strongly than by the usual subformula property.

However, pure forward proof search techniques such as the inverse method also have shortcomings. For instance, the existence of two  $\triangleright R$  rules is a concession to the need for localized weakening, something usually handled more elegantly in backward decision procedures by general weakening. Also, the refined focusing we have introduced strongly restricts what rules are applicable, something that a decision procedure should be able to exploit in order to generate fewer intermediate sequents. These issues are addressed in the next section by combining ideas from forward and backward proof search.

## 4 Bidirectional Proof Search in Sequent Calculus

The idea behind the bidirectional sequent calculus method is that given a query formula  $A$ , we can, by exploiting forward proof search techniques, but before performing proof search itself, construct a set of derived inference rules for  $\mathbf{MJ}_{\mathbf{IS4}}$  which conceal all left-rule applications that could be needed in a proof of  $A$ . We then carry out backward proof search over these relevant derived rules and the usual right-rules of  $\mathbf{MJ}_{\mathbf{IS4}}$ . By design, our derived inference rules will correspond exactly to the notion of *focused threads* in  $\mathbf{MJ}_{\mathbf{IS4}}^F$  derivations, defined as follows.

**Definition 2 (Focused threads).** *A focused thread of an  $\mathbf{MJ}_{\mathbf{IS4}}^F$  derivation is a segment of the derivation that begins, at the top, with an application of  $\text{init}_i$ ,  $\perp L_i$ ,  $\vee L_i$ ,  $\square L_i$ , or  $\diamond L_i$ ,  $i \in \{1, 2\}$  (raising a formula into a stoup), includes only focused sequents, and ends with an application of  $\text{ch}_i$  (dropping a formula from the stoup).*

In any  $\mathbf{MJ}_{\mathbf{IS}_4}^{\mathbf{F}}$  derivation of an unfocused sequent, left-rule applications must occur in focused threads, so we can think of derivations as consisting of focused threads strung together using right-rule applications. The key insight is that all focused threads possibly needed in an  $\mathbf{MJ}_{\mathbf{IS}_4}^{\mathbf{F}}$  proof of a formula  $A$  can be deterministically constructed prior to proof search by inspecting the structure of  $A$ . To justify this claim, we will use our refined subformula property.

First note that it is straightforward to uniquely label subformula occurrences of a formula to be proved, and that the definition of signed subformulas, the signed subformula property, and the inference rules of  $\mathbf{MJ}_{\mathbf{IS}_4}^{\mathbf{F}}$  can be adjusted to operate on labels rather than formulas, thus differentiating between subformula occurrences.

To give some intuition as to how to construct all the focused threads possibly needed for a proof of a formula, we will illustrate the approach on the following small example:

$$\begin{array}{c}
 \overbrace{\hspace{10em}}^{L_0^+} \\
 \overbrace{\hspace{4em}}^{L_1^-, L_1^\sim} \quad \overbrace{\hspace{4em}}^{L_5^+} \\
 \overbrace{\hspace{4em}}^{L_2^-, L_2^\approx} \quad \overbrace{\hspace{4em}}^{L_6^+} \\
 \overbrace{\hspace{2em}}^{L_3^+} \quad \overbrace{\hspace{2em}}^{L_4^\approx} \quad \overbrace{\hspace{2em}}^{L_7^-, L_7^\sim} \quad \overbrace{\hspace{2em}}^{L_8^+} \\
 \square(\overbrace{A}^+ \supset \overbrace{B}^\approx) \supset \diamond(\overbrace{A}^- \supset \overbrace{B}^+)
 \end{array}$$

with subformulas

$$L_0^+, L_3^+, L_5^+, L_6^+, L_8^+, \quad L_1^-, L_7^-, \quad L_1^\sim, L_7^\sim, \quad L_2^-, \quad \text{and} \quad L_2^\approx, L_4^\approx.$$

The signed subformula property guarantees that in a proof of the sequent  $\cdot; \cdot \mapsto L_0^+$ , the only axioms we require are

$$\frac{}{\cdot; \cdot \triangleright L_7^\sim \mapsto L_3^+} \text{init}_1 \quad \text{and} \quad \frac{}{\cdot; \cdot \triangleright \triangleright L_4^\approx \mapsto L_8^+} \text{init}_2$$

Consider the first of these axioms. Every left-rule either drops the stoup formula into a context or expands it. The immediate parent of  $L_7^\sim$  in the subformula hierarchy is  $L_7^-$ , indicating that dropping  $L_7$  into the context is a permissible operation. In fact, it is the *only* operation permitted by the signed subformula property operating on labels. We can collapse this short focused thread into a single derived inference rule:

$$\frac{\frac{}{\cdot; \cdot \triangleright L_7^\sim \mapsto L_3^+} \text{init}_1}{\cdot; L_7^- \mapsto L_3^+} \text{ch}_1}{\cdot; L_7^- \mapsto L_3^+} \sim \frac{}{\cdot; L_7^- \mapsto L_3^+} (1)$$

Considering the second axiom, we notice that the parent subformula of  $L_4^\approx$  is  $L_2^\approx$ , also a focused subformula. The next rule application should then be  $\triangleright L_2$ , with  $L_2^\approx$  as the principal formula. In fact, it is not difficult to see that since every subformula occurrence has a unique parent subformula, the signed subformula property operating on labels always uniquely dictates which rule may be applied. This game continues until the end of the focused thread. In the case of the second

axiom, the immediate parent of  $L_2^\approx$  is  $L_2^-$ , signalling an application of  $\text{ch}_2$  and the end of the thread:

$$\frac{\frac{\Delta; \Gamma \mapsto L_3^+ \quad \overline{\cdot; \triangleright L_4^\approx \mapsto L_8^+} \text{init}_2}{\Delta; \Gamma \triangleright L_2^\approx \mapsto L_8^+} \supset L_2}{\Delta, L_2^-; \Gamma \mapsto L_8^+} \text{ch}_2 \quad \rightsquigarrow \quad \frac{\Delta; \Gamma \mapsto L_3^+}{\Delta, L_2^-; \Gamma \mapsto L_8^+} \quad (2)$$

Note that this thread, unlike the one concealed by (1), has open premises and is parametric in the contexts  $\Delta$  and  $\Gamma$ . Finally, the signed subformula property allows one more focused thread, starting with

$$\frac{\Delta, L_2^-; \Gamma \mapsto M^+}{\Delta; \Gamma \triangleright L_1^\sim \mapsto M^+} \Box L_1$$

The immediate parent subformula of  $L_1^\sim$  is  $L_1^-$ , so this thread ends here, yielding the derived rule

$$\frac{\frac{\Delta, L_2^-; \Gamma \mapsto M^+}{\Delta; \Gamma \triangleright L_1^\sim \mapsto M^+} \Box L_1}{\Delta; \Gamma, L_1^- \mapsto M^+} \text{ch}_1 \quad \rightsquigarrow \quad \frac{\Delta, L_2^-; \Gamma \mapsto M^+}{\Delta; \Gamma, L_1^- \mapsto M^+} \quad (3)$$

Notice that this big step rule is schematic not only in the contexts  $\Delta$  and  $\Gamma$ , but also in the goal formula  $M^+$ . Since the signed subformula property allows no other focused threads, the remainder of the proof, if one exists, may only chain the derived rules (1), (2), and (3) together with right-rule applications. In this case, completing the proof is straightforward:

$$\begin{array}{c} \overline{\cdot; L_7^- \mapsto L_3^+} \quad (1) \\ \frac{\cdot; L_7^- \mapsto L_3^+}{L_2^-; L_7^- \mapsto L_8^+} \quad (2) \\ \supset R \\ \frac{L_2^-; \cdot \mapsto L_6^+}{L_2^-; \cdot \mapsto L_5^+} \quad \diamond R \\ \frac{L_2^-; \cdot \mapsto L_5^+}{\cdot; L_1^- \mapsto L_5^+} \quad (3) \\ \supset R \\ \cdot; \cdot \mapsto L_0^+ \end{array}$$

In general, to cover all focused threads, the construction of derived rules must begin with focused sequents of the following kinds, where  $*$  is  $\sim$  or  $\approx$ , depending on whether  $i = 1$  or  $i = 2$ :

1.  $\cdot; \triangleright^i L_j^* \mapsto L_k^+$ , where  $L_j$  and  $L_k$  denote the same atomic formula,
2.  $\cdot; \triangleright^i L_j^* \mapsto M^+$ , where  $L_j$  denotes  $\perp$  and  $M$  is schematic,
3.  $\Delta; \Gamma \triangleright^i L_j^* \mapsto M^+$ , where  $L_j$  denotes some  $A_1 \vee A_2$  and  $M$  is schematic,
4.  $\Delta; \Gamma \triangleright^i L_j^* \mapsto M^+$ , where  $L_j$  denotes some  $\Box A$  and  $M$  is schematic, and
5.  $\Delta; \Gamma \triangleright^i L_j^* \mapsto M^+$ , where  $L_j$  denotes some  $\diamond A$  and  $M$  denotes some  $\diamond C$ ,  $C$  being schematic.

Moreover, the constructed derived rules must end with a stoup formula being dropped into one of the contexts.

The question now is how these forward-constructed derived rules can complement backward proof search. The key observation is that every focused thread of an  $\mathbf{MJ}_{\mathbf{IS4}}^{\mathbf{F}}$  derivation can be converted into a focused thread of an  $\mathbf{MJ}_{\mathbf{IS4}}$  derivation by applying weakening, reducing valid focused sequents to focused sequents, and omitting the now unnecessary signs of subformula labels. For instance,

$$\frac{\frac{\Delta; \Gamma \mapsto L_3^+ \quad \overline{\cdot \triangleright \triangleright L_4^{\approx} \mapsto L_8^+}}{\Delta; \Gamma \triangleright \triangleright L_2^{\approx} \mapsto L_8^+} \text{init}_2}{\Delta, L_2^-; \Gamma \mapsto L_8^+} \triangleright L_2 \text{ch}_2$$

can be converted into the  $\mathbf{MJ}_{\mathbf{IS4}}$  derivation segment

$$\frac{\frac{\Delta_1, L_2, \Delta_2; \Gamma \rightarrow L_3 \quad \overline{\Delta_1, L_2, \Delta_2; \Gamma \triangleright L_4 \rightarrow L_8}}{\Delta_1, L_2, \Delta_2; \Gamma \triangleright L_2 \rightarrow L_8} \text{init}}{\Delta_1, L_2, \Delta_2; \Gamma \rightarrow L_8} \triangleright L \text{ch}_2$$

This makes it possible to construct derived rules for  $\mathbf{MJ}_{\mathbf{IS4}}$ . The benefit of performing backward proof search over these derived rules and the remaining right rules is that it requires no conventional loop-detection. However, some bookkeeping is still required, since our bidirectional decision procedure has one important termination requirement: that every derived rule instance — characterized by the identity of the schematic derived rule and the concrete goal formula, if applicable — is used at most once along every branch of the proof, from root to leaf. The following result of  $\mathbf{MJ}_{\mathbf{IS4}}$  guarantees that this requirement does not cost us completeness.

**Theorem 4 (Uniqueness of stoup and goal formula occurrences on branches).** *If a sequent  $\Delta; \Gamma \rightarrow L$  is derivable, then it has a derivation with the property that no branch (from root to leaf) contains more than one application of  $\text{ch}_i$ ,  $i \in \{1, 2\}$  with the same stoup and goal formula occurrences.*

*Proof.* A derivation with loops of this kind can be shortened by collapsing segments between repeated applications of  $\text{ch}_i$ ,  $i \in \{1, 2\}$ .  $\square$

Since the identity of a focused thread depends on the identities of the focused formula occurrences it contains, and on its goal formula occurrence, we obtain the following important corollary.

**Corollary 1.** *If a sequent  $\Delta; \Gamma \rightarrow L$  is derivable, then it has a derivation with the property that no focused thread instance occurs more than once along a branch.*

The consequence of this result is that if a sequent is provable in  $\mathbf{MJ}_{\mathbf{IS4}}$ , then it is provable without using any derived rule instance more than once along

a branch. With the observation that every right rule of  $\mathbf{MJ}_{\mathbf{IS4}}$  reduces the complexity of the goal formula, this means that every rule application during backward proof search in  $\mathbf{MJ}_{\mathbf{IS4}}$  with derived rules either reduces the number of available derived rule instances along the current branch, or leaves the number of available derived rule instances unchanged but reduces the complexity of the goal formula. This measure gives an immediate termination guarantee without the need for conventional loop-detection. All that is needed is a way of keeping track of which derived rule instances have been applied along a branch. While this bookkeeping apparatus is reminiscent of a history mechanism, we expect it to be far more lightweight than maintaining histories of previously encountered sequents or goal formulas, as is common in standard loop-detection schemes. Our expectations will, for the most part, be vindicated by our experimental results.

Note that the idea of constructing relevant derived rules prior to proof search can also be exploited in forward proof search, where the derived rules described above can take the place of left rules in the inverse method. The main advantages here are that the derived rules are more relevant to proof search for the given query formula, and that the number of intermediate sequents added to the knowledge base during proof search is reduced, since no focused sequents need to be maintained.

## 5 Bidirectional Proof Search in Natural Deduction

In the backward bidirectional sequent calculus method, we construct derived rules to conceal all required focused threads. Notice that the focused threads of  $\mathbf{MJ}_{\mathbf{IS4}}^{\mathbf{F}}$  correspond naturally to segments of  $\mathbf{NJ}_{\mathbf{IS4}}^{\mathbf{N}}$  proofs consisting of elimination rule applications, that is,  $\downarrow$  judgements. The beginnings of focused threads, where formulas are placed into the stoup, correspond to *reversing rules* in  $\mathbf{NJ}_{\mathbf{IS4}}^{\mathbf{N}}$ . These are the  $\uparrow\downarrow$  rule, as well as all elimination rules with  $\uparrow$  judgements as their conclusions. The ends of focused threads, on the other hand, where the stoup formula is dropped into a context, correspond to using a hypothesis with applications of  $\text{hyp}_1$  or  $\text{hyp}_2$ .

This means that the process of building a derived  $\mathbf{MJ}_{\mathbf{IS4}}^{\mathbf{F}}$  rule in a top-down way corresponds to building a natural deduction derived rule by beginning with an application of a reversing rule, and growing it upwards until we reach a leaf. Just as the construction of derived rules in the sequent calculus is determined uniquely by the form of the query formula, so these natural deduction derived rules can be deterministically constructed before proof search even begins.

This approach is best demonstrated by an example such as the one given in Sect. 4. For instance, given the pair  $L_4^{\approx}$  and  $L_8^+$  from that example, we begin with the coercion

$$\frac{\Delta; \Gamma \vdash L_4^{\approx} \downarrow}{\Delta; \Gamma \vdash L_8^+ \uparrow} \uparrow\downarrow$$

Since the immediate parent of  $L_4^{\approx}$  in the signed subformula hierarchy is  $L_2^{\approx}$ , denoting  $A \supset B$ , the rule application above this coercion must be an application

of  $\supset E$ :

$$\frac{\frac{\Delta; \Gamma \vdash L_2^{\approx} \downarrow \quad \Delta; \Gamma \vdash L_3^+ \uparrow}{\Delta; \Gamma \vdash L_4^{\approx} \downarrow} \supset E}{\Delta; \Gamma \vdash L_8^+ \uparrow} \updownarrow$$

The focused thread continues along the first premise, but the parent of  $L_2^{\approx}$  is  $L_2^{\bar{}}$ , indicating the end of this focused thread by an application of  $\text{hyp}_2$ :

$$\frac{\frac{\frac{L_2^{\bar{}} \in \Delta}{\Delta; \Gamma \vdash L_2^{\approx} \downarrow} \text{hyp}_2 \quad \Delta; \Gamma \vdash L_3^+ \uparrow}{\Delta; \Gamma \vdash L_4^{\approx} \downarrow} \supset E}{\Delta; \Gamma \vdash L_8^+ \uparrow} \updownarrow}{\Delta_1, L_2^{\bar{}}, \Delta_2; \Gamma \vdash L_3^+ \uparrow} \sim \frac{\Delta_1, L_2^{\bar{}}, \Delta_2; \Gamma \vdash L_3^+ \uparrow}{\Delta_1, L_2^{\bar{}}, \Delta_2; \Gamma \vdash L_8^+ \uparrow} \quad (2)$$

In similar constructions, the pair  $L_7^{\approx}$ ,  $L_3^+$  and  $L_1^{\sim}$ , the latter denoting  $\Box(A \supset B)$ , produce, respectively, the natural deduction derived rules

$$\frac{\frac{L_7^{\bar{}} \in \Gamma}{\Delta; \Gamma \vdash L_7^{\approx} \downarrow} \text{hyp}_1}{\Delta; \Gamma \vdash L_3^+ \uparrow} \updownarrow \sim \frac{}{\Delta; \Gamma_1, L_7^{\bar{}}, \Gamma_2 \vdash L_3^+ \uparrow} \quad (1)$$

and

$$\frac{\frac{\frac{L_1^{\bar{}} \in \Gamma}{\Delta; \Gamma \vdash L_1^{\sim} \downarrow} \text{hyp}_1 \quad \Delta, L_2^{\bar{}}; \Gamma \vdash M^+ \uparrow}{\Delta; \Gamma \vdash M^+ \uparrow} \Box E}{\Delta; \Gamma \vdash M^+ \uparrow} \sim \frac{\Delta, L_2^{\bar{}}; \Gamma_1, L_1^{\bar{}}, \Gamma_2 \vdash M^+ \uparrow}{\Delta; \Gamma_1, L_1^{\bar{}}, \Gamma_2 \vdash M^+ \uparrow} \quad (3)$$

The rest of the proof then uses only these derived rules and introduction rules:

$$\begin{array}{l} \frac{}{L_2^{\bar{}}; L_1^{\bar{}}, L_7^{\bar{}} \vdash L_3^+ \uparrow} \quad (1) \\ \frac{}{L_2^{\bar{}}; L_1^{\bar{}}, L_7^{\bar{}} \vdash L_8^+ \uparrow} \quad (2) \\ \frac{}{L_2^{\bar{}}; L_1^{\bar{}} \vdash L_6^+ \uparrow} \supset I \\ \frac{}{L_2^{\bar{}}; L_1^{\bar{}} \vdash L_5^+ \uparrow} \diamond I \\ \frac{}{; L_1^{\bar{}} \vdash L_5^+ \uparrow} \quad (3) \\ \frac{}{; \cdot \vdash L_0^+ \uparrow} \supset I \end{array}$$

In general, the approach for constructing natural deduction derived rules is analogous to the method for the backward bidirectional sequent calculus, only turned upside-down, in the sense that the rule at the beginning of an  $\mathbf{MJ}_{\text{IS4}}^F$  focused thread determines the reversing rule at the bottom of the natural deduction focused thread, while the final application of  $\text{hyp}_i$  dictates the ‘‘principal formula’’ of the ensuing derived natural deduction rule.

Proof search over natural deductions can then be performed in a backward direction. The only nondeterminism is in whether to apply a derived rule or an introduction rule, the premises of which are uniquely determined by their conclusions. Note that to guarantee termination, we again disallow using a derived rule instance more than once along any branch of a proof.

**Table 1.** Selection of experimental results

Formula Size	Modalities	Provable	Histories Time	Inverse Time	Rules	Bidirectional Time	Rules	
32	49	0	N	> 1000	1.36	33	0.01	33
36	175	0	Y	0.08	> 1000	159	> 1000	592
37	68	9	Y	84.79	1.18	60	< 0.01	28
39	42	3	N	8.46	1.83	31	< 0.01	15
44	49	14	Y	75.13	> 1000	51	37.11	21
50	44	7	Y	7.38	> 1000	49	48.76	25

## 6 Experimental Results

While benchmark formulas are available for intuitionistic propositional logic and classical modal logics, we are not aware of any benchmark libraries specific to propositional **IS4**. In order to evaluate the performance of our bidirectional approach, we put together a benchmark set of 50 formulas for **IS4**, mostly problems from Rathes et al.’s Intuitionistic Logic Theorem Proving (ILTP) library [17] to which we introduced modalities. Our full benchmark set is provided as an appendix to the accompanying technical report [10].

We implemented three **IS4** decision procedures in SML: (1) an  $\mathbf{MJ}_{\mathbf{IS4}}$ -based backward prover with a history mechanism for loop-detection, (2) an  $\mathbf{MJ}_{\mathbf{IS4}}^{\mathbf{F}}$ -based inverse method prover without derived rules, and (3) our bidirectional natural deduction prover. The loop-detection prover maintains two histories to detect repeated modal and nonmodal rule applications, respectively. This approach is a generalization of Howe’s decision procedure [11] extended to full **IS4**. Note that the behaviour of our backward bidirectional sequent calculus prover corresponds exactly to that of the bidirectional natural deduction prover, so we have only implemented the more elegant natural deduction prover.

On many of the smaller problems, there was little measurable difference in the performance of the provers, but some of the problems that did elicit noticeably different performances are highlighted in Table 1. The size column shows the complexity of each formula, computed inductively in the usual way, while the modalities column shows the number of modal operators. Times are in seconds.<sup>1</sup> For the inverse method and bidirectional provers, we show the number of inference rules generated (derived rules in the case of the bidirectional prover).

As the results demonstrate, the bidirectional natural deduction prover is a competitive alternative to the more conventional provers, equalling or outperforming them on most problems. Comparing the average proving time for problems that were solved, it is noticeably superior, although we found two formulas on which it was significantly outperformed (formulas 36 and 50 in Table 1). Interestingly, there is not always a clear connection between the number of derived

<sup>1</sup> All timing results were obtained on a Pentium III 850 MHz with 256 MB of RAM, running SML/NJ version 110.60.

rules generated and the time required to solve a problem. Presumably, the problematic cases were those whose derived rules were the shortest and least useful. Note also that as derived rules are associated with subformula occurrences, those formulas with many repeated subformulas (e.g. formula 36) caused a very large number of duplicate derived rules to be generated.

## 7 Related and Future Work

Although **IS4** has undergone thorough proof-theoretical studies, there has been little work in developing proof search strategies specific to it. We have presented a comprehensive study of proof search formalisms for **IS4**, highlighting the duality between backward and forward search. Moreover, we have demonstrated how to combine the benefits of both to yield bidirectional decision procedures based on sequent calculi and natural deduction. Our experimental results reveal that combining the two traditionally disparate paradigms can be fruitful. Although our implementations are naive and unoptimized, we hope that our results might encourage further study of bidirectional proof search, particularly in other logics.

For instance, in the contextual modal logic of Nanevski, Pfenning, and Pientka [14], structural modality is generalized by relativizing the validity judgement and the modal operators. The techniques discussed in this paper extend very naturally to contextual modal logic, yielding sequent calculi suitable for backward and forward proof search, but the exact nature of how such a generalization affects proof search is yet to be explored. The reconciliation of forward and backward proof search has recently also been investigated by Chaudhuri and Pfenning [2], who, in the context of linear logic, propose a focusing inverse method prover incorporating derived rules constructed in a backward way and searched over in a forward direction, precisely opposite to our approach.

In the future, we plan to explore extensions to the first-order case. Although the idea of derived rules extends, in principle, to first-order quantifiers, the constructed derived rules become parametric in terms. The useful property of **MJIS4** that eliminated the need for conventional loop-detection in our bidirectional method now only holds for particular instantiations of the terms of the parametric derived rules. Unfortunately, requiring the storage of rule instantiations introduces another layer of bookkeeping. How to efficiently overcome this problem and what the proof-theoretical relationship between first-order bidirectional decision procedures and natural deduction provers is remains to be investigated.

**Acknowledgements.** We would like to thank Daniel Pomerantz for discussions on forward proof search and for the implementation of the inverse method prover, and the reviewers for their insightful comments and suggestions.

## References

1. G. M. Bierman and V. C. V. de Paiva. On an intuitionistic modal logic. *Studia Logica*, 65(3):383–416, 2000.

2. K. Chaudhuri and F. Pfenning. Focusing the inverse method for linear logic. In C.-H. L. Ong (ed.) *Proceedings of the 19th International Workshop on Computer Science Logic, 14th Annual Conference of the EACSL (CSL'05)*, pp. 200–215. Springer LNCS 3634, 2005.
3. R. Davies and F. Pfenning. A modal analysis of staged computation. *Journal of the ACM*, 48(3):555–604, 2001.
4. A. Degtyarev and A. Voronkov. The inverse method. In J. A. Robinson and A. Voronkov (eds.) *Handbook of Automated Reasoning*, pp. 179–272. Elsevier and MIT Press, 2001.
5. R. Dyckhoff and L. Pinto. A permutation-free sequent calculus for intuitionistic logic. Research Report CS/96/9, University of St. Andrews, 1996.
6. M. Fairtlough and M. Mendler. Propositional lax logic. *Information and Computation*, 137(1):1–33, 1997.
7. F. B. Fitch. Intuitionistic modal logic with quantifiers. *Portugaliae Mathematica*, 7(2):113–118, 1948.
8. J.-Y. Girard. A new constructive logic: Classical logic. *Mathematical Structures in Computer Science*, 1(3):255–296, 1991.
9. J.-Y. Girard, Y. Lafont, and P. Taylor. *Proofs and types*. Cambridge University Press, 1989.
10. S. Heilala and B. Pientka. Bidirectional decision procedures for the intuitionistic propositional modal logic IS4. Technical Report SOCS-TR-2007.2, McGill University, May 2007.
11. J. M. Howe. Proof search issues in some non-classical logics. PhD thesis, University of St. Andrews, 1998.
12. J. M. Howe. Proof search in lax logic. *Mathematical Structures in Computer Science*, 11(4): 573–588, 2001.
13. J. Moody. Modal logic as a basis for distributed computation. Technical Report CMU-CS-03-194, Carnegie Mellon University, 2003.
14. A. Nanevski, F. Pfenning, and B. Pientka. Contextual modal type theory. To appear in *ACM Transactions on Computational Logic*.
15. F. Pfenning and R. Davies. A judgmental reconstruction of modal logic. *Mathematical Structures in Computer Science*, 11(4):511–540, 2001.
16. D. Prawitz. *Natural deduction: A proof-theoretical study*. Almqvist & Wiksell, 1965.
17. T. Raths, J. Otten, and C. Kreitz. The ILTP problem library for intuitionistic logic, release v1.1. To appear in *Journal of Automated Reasoning*.
18. T. W. Satre. Natural deduction rules for modal logics. *Notre Dame Journal of Formal Logic*, 13(4):461–475, 1972.
19. A. K. Simpson. The proof theory and semantics of intuitionistic modal logic. PhD thesis, University of Edinburgh, 1994.