Focusing the Inverse Method for Linear Logic

Kaustuv Chaudhuri and Frank Pfenning*

Department of Computer Science
Carnegie Mellon University
kaustuv@cs.cmu.edu and fp@cs.cmu.edu

Abstract. Focusing is traditionally seen as a means of reducing inessential non-determinism in backward-reasoning strategies such as uniform proof-search or tableaux systems. In this paper we construct a form of focused derivations for propositional linear logic that is appropriate for forward reasoning in the inverse method. We show that the focused inverse method conservatively generalizes the classical hyperresolution strategy for Horn-theories, and demonstrate through a practical implementation that the focused inverse method is considerably faster than the non-focused version.

1 Introduction

Strategies for automated deduction can be broadly classified as backward reasoning or forward reasoning. Among the backward reasoning strategies we find tableaux and matrix methods; forward reasoning strategies include resolution and the inverse method. The approaches seem fundamentally difficult to reconcile because the state of a backward reasoner is global, while a forward reasoner maintains locally self-contained state.

Both backward and forward approaches are amenable to reasoning in non-classical logics. This is because they can be derived from an inference system that defines a logic. The derivation process is systematic to some extent, but in order to obtain an effective calculus and an efficient implementation, we need to analyze and exploit deep proof-theoretic or semantic properties of each logic under consideration.

Some themes stretch across both backwards and forwards systems and even different logics. Cut-elimination and its associated subformula property, for example, are absolutely fundamental for both types of systems, regardless of the underlying logic. In this paper we advance the thesis that *focusing* is similarly universal. Focusing was originally designed by Andreoli [1, 2] to remove inessential non-determinism from backward proof search in classical linear logic. It has already been demonstrated [3] that focusing applies to other logics; here we show that focusing is an important concept for theorem proving in the forward direction.

As the subject of our study we pick propositional intuitionistic linear logic [4–6].This choice is motivated by two considerations. First, it includes the propositional core of the Concurrent Logical Framework (CLF), so our theorem prover, and its first-order extension, can reason with specifications written in CLF; many such specifications, including Petri nets, the π-calculus and Concurrent ML, are described in [7].

* This work has been supported by the Office of Naval Research (ONR) under grant MURI N00014-04-1-0724 and by the National Science Foundation (NSF) under grant CCR-0306313.

For many of these applications, the intuitionistic nature of the framework is essential. Second, it is almost a worst-case scenario, combining the difficulties of modal logic, intuitionistic logic, and linear logic, where even the propositional fragment is undecidable. A treatment, for example, of classical linear logic without the lax modality can be given very much along the same lines, but would be simpler in several respects.

Our contributions are as follows. First, we show how to construct a non-focusing inverse method for intuitionistic linear logic. This follows a fairly standard recipe [8], although the resource management problem germane to linear logic has to be considered carefully. Second, we define focused derivations for intuitionistic linear logic. The focusing properties of the connectives turn out to be consistent with their classical interpretation, but completeness does not come for free because of the additional restrictions placed by intuitionistic (and modal) reasoning. The completeness proof is also somewhat different from ones we have found in the literature. Third, we show how to adapt focusing so it can be used in the inverse method. The idea is quite general and, we believe, can be adapted to other non-classical logics. Fourth, we demonstrate via experimental results that the focused inverse method is substantially faster than the non-focused one. Fifth, we show that refining the inverse method with focusing agrees exactly with classical hyperresolution on Horn formulas, a property which fails for non-focusing versions of the inverse method. This is practically significant, because even in the linear setting many problems or subproblems may be non-linear and Horn, and need to be treated with reasonable efficiency.

In a related paper [9] we generalize our central results to first-order intuitionistic linear logic, provide more detail on the implementation choices, and give a more thorough experimental evaluation. Lifting the inverse method here to include quantification is far from straightforward, principally because of the rich interactions between linearity, weakening, and contraction in the presence of free variables. However, these considerations are orthogonal to the basic design of forward focusing which remains unchanged from the present paper.

Perhaps most closely related to our work is Tammet's inverse method prover for classical linear logic [10] which is a refinement of Mints' resolution system [11]. Some of Tammet's optimizations are similar in nature to focusing, but are motivated primarily by operational rather than by logical considerations. As a result, they are not nearly as far-reaching, as evidenced by the substantial speedups we obtain with respect to Tammet's implementation. Our examples were chosen so that the difference between intuitionistic and classical linear reasoning was inessential.

2 Backward linear sequent calculus

We use a backward cut-free sequent calculus for propositions constructed out of the propositional linear connectives $\{\otimes, \mathbf{1}, \multimap, \&, \top, !\}$; the extension to first-order connectives using the recipe outlined in [9] is straightforward. To simplify the presentation we leave out $\oplus$ and $\mathbf{0}$, though the implementation supports them and some of the experiments in Sec. 5.2 use them. Propositions are written using uppercase letters A, B, C, with p standing for atomic propositions. The sequent calculus is a standard fragment of JILL [6], containing dyadic two-sided sequents of the form $\Gamma ; \Delta \Longrightarrow C$: the zone Γ

contains the unrestricted hypotheses and Δ contains the linear hypotheses. Both contexts are unordered. For the rules of this calculus we refer the reader to [6, page 14]. Also in [6] are the standard weakening and contraction properties for the unrestricted hypotheses, which means we can treat Γ as a set, and admissibility of cut by means of a simple lexicographic induction.

Definition 1 (subformulas). *A* decorated formula *is a tuple $\langle A, s, w\rangle$ where A is a proposition, s is a* sign *(+ or −) and w is a* weight *(h for* heavy *or l for* light*). The* subformula relation $\leq$ *is the smallest reflexive and transitive relation between decorated subformulas satisfying the following inequalities:*

$$\langle A, s, h\rangle \leq \langle ! A, s, *\rangle \quad \langle A, \overline{s}, l\rangle \leq \langle A \multimap B, s, *\rangle \quad \langle B, s, l\rangle \leq \langle A \multimap B, s, *\rangle$$
$$\langle A_i, s, l\rangle \leq \langle A_1 \otimes A_2, s, *\rangle \quad \langle A_i, s, l\rangle \leq \langle A_1 \,\&\, A_2, s, *\rangle \qquad i \in \{1, 2\}$$

where $\overline{s}$ is the opposite of s, and $*$ can stand for either h or l, as necessary. Decorations and the subformula relation are lifted to (multi)sets in the obvious way.

Property 2 (subformula property). *In any sequent $\Gamma' ; \Delta' \Longrightarrow C'$ used in a proof of $\Gamma ; \Delta \Longrightarrow C$: $\langle \Gamma', -, h\rangle \cup \langle \Delta', -, *\rangle \cup \{\langle C', +, *\rangle\} \leq \langle \Gamma, -, h\rangle \cup \langle \Delta, -, l\rangle \cup \{\langle C, +, l\rangle\}$.* □

For the remainder of the paper, all rules are restricted to decorated subformulas of a given goal sequent. A right (resp. left) rule is applicable if the principal formula in the conclusion is a positive (resp. negative) subformula of the goal sequent. Of the judgmental rules (reviewed in the next section), init is restricted to atomic subformulas that are *both* positive and negative decorated subformulas, and the copy rule is restricted to negative heavy subformulas.

3 Forward linear sequent calculus

In addition to the usual non-determinism in rule and subgoal selection, the single-use semantics of linear hypotheses gives rise to *resource non-determinism* during backward search. Its simplest form is *multiplicative*, caused by binary multiplicative rules ($\otimes R$ and $\multimap L$), where the linear zone of the conclusion has to be distributed into the premisses. In order to avoid an exponential explosion, backward search strategies postpone this split either by an input/output interpretation, where proving a subgoal consumes some of the resources from the input and passes the remaining resources on as outputs [12], or via Boolean constraints on the occurrences of linear hypotheses [13]. Interestingly, multiplicative non-determinism is entirely absent in a forward reading of multiplicative rules: the linear context in the conclusion is formed simply by adjoining those of the premisses. On the multiplicative-exponential fragment, for example, forward search has no resource management issues at all. Resource management problems remain absent even in the presence of binary additives (& and ⊕).

The only form of resource non-determinism for the forward direction arises in the presence of additive constants ($\top$ and $\mathbf{0}$). For example, the backward $\top R$ rule has an arbitrary linear context which we cannot guess in the forward direction. We therefore leave it empty (no linear assumptions are needed), but we have to remember that we can

judgmental

$$\frac{}{\cdot \,;\, p \longrightarrow^{0} p}\ \text{init} \qquad \frac{\Gamma \,;\, \Delta, A \longrightarrow^{w} C}{\Gamma \cup \{A\} \,;\, \Delta \longrightarrow^{w} C}\ \text{copy}$$

multiplicative

$$\frac{\Gamma \,;\, \Delta \longrightarrow^{w} A \quad \Gamma' \,;\, \Delta' \longrightarrow^{w'} B}{\Gamma \cup \Gamma' \,;\, \Delta, \Delta' \longrightarrow^{w \vee w'} A \otimes B}\ \otimes R$$

$$\frac{\Gamma \,;\, \Delta, A, B \longrightarrow^{w} C}{\Gamma \,;\, \Delta, A \otimes B \longrightarrow^{w} C}\ \otimes L$$

$$\frac{\Gamma \,;\, \Delta, A_i \longrightarrow^{1} C \quad (A_j \notin \Delta)}{\Gamma \,;\, \Delta, A_1 \otimes A_2 \longrightarrow^{1} C}\ \otimes L_i \qquad (i, j) \in \{(1, 2), (2, 1)\}$$

$$\frac{}{\cdot \,;\, \cdot \longrightarrow^{0} \mathbf{1}}\ \mathbf{1}R \qquad \frac{\Gamma \,;\, \Delta \longrightarrow^{0} C}{\Gamma \,;\, \Delta, \mathbf{1} \longrightarrow^{0} C}\ \mathbf{1}L$$

$$\frac{\Gamma \,;\, \Delta, A \longrightarrow^{w} B}{\Gamma \,;\, \Delta \longrightarrow^{w} A \multimap B}\ \multimap R$$

$$\frac{\Gamma \,;\, \Delta \longrightarrow^{1} B \quad (A \notin \Delta)}{\Gamma \,;\, \Delta \longrightarrow^{1} A \multimap B}\ \multimap R'$$

$$\frac{\Gamma \,;\, \Delta, B \longrightarrow^{w} C \quad \Gamma' \,;\, \Delta' \longrightarrow^{w'} A \quad (w = 0 \vee B \notin \Delta')}{\Gamma \cup \Gamma' \,;\, \Delta, \Delta', A \multimap B \longrightarrow^{w \vee w'} C}\ \multimap L$$

additive

$$\frac{\Gamma \,;\, \Delta \longrightarrow^{w} A \quad \Gamma' \,;\, \Delta' \longrightarrow^{w'} B \quad (\Delta / w \approx \Delta' / w')}{\Gamma \cup \Gamma' \,;\, \Delta \sqcup \Delta' \longrightarrow^{w \wedge w'} A \,\&\, B}\ \&R$$

$$\frac{}{\cdot \,;\, \cdot \longrightarrow^{1} \top}\ \top R \qquad \frac{\Gamma \,;\, \Delta, A_i \longrightarrow^{w} C}{\Gamma \,;\, \Delta, A_1 \,\&\, A_2 \longrightarrow^{w} C}\ \&L_i \qquad i \in \{1, 2\}$$

exponential

$$\frac{\Gamma \,;\, \cdot \longrightarrow^{w} A}{\Gamma \,;\, \cdot \longrightarrow^{0} !A}\ !R \qquad \frac{\Gamma, A \,;\, \Delta \longrightarrow^{w} C}{\Gamma \,;\, \Delta, !A \longrightarrow^{w} C}\ !L$$

$$\frac{\Gamma \,;\, \Delta \longrightarrow^{0} C \quad (A \notin \Gamma)}{\Gamma \,;\, \Delta, !A \longrightarrow^{0} C}\ !L'$$

Fig. 1: forward linear sequent calculus

add linear assumptions if necessary. We therefore differentiate sequents whose linear context can be weakened and those whose can not.

To distinguish forward from backward sequents, we shall use a single arrow ($\longrightarrow$), possibly decorated, but keep the names of the rules the same.

Definition 3 (forward sequents).

1. *A* forward sequent *is of the form $\Gamma \,;\, \Delta \longrightarrow^{w} C$. Γ and Δ hold the unrestricted and linear resources respectively, and w is a Boolean (0 or 1) called the* weak-flag. *Sequents with $w = 1$ are called* weakly linear *or simply* weak, *and those with $w = 0$ are* strongly linear *or* strong.
2. *The* correspondence relation $\prec$ *between forward and backward sequents is defined as follows: $(\Gamma \,;\, \Delta \longrightarrow^{w} C) \prec (\Gamma' \,;\, \Delta' \Longrightarrow C)$ iff $\Gamma \subseteq \Gamma'$, and $\Delta = \Delta'$ or $\Delta \subseteq \Delta'$ depending on whether $w = 0$ or $w = 1$, respectively. The forward sequent s is* sound *if for every backward sequent s' such that $s \prec s'$, s' is derivable in the backward calculus.*
3. *The* subsumption relation $\leq$ *between forward sequents is the smallest relation to satisfy:*

$$\left.\begin{array}{l}(\Gamma \,;\, \Delta \longrightarrow^{0} C) \leq (\Gamma' \,;\, \Delta \longrightarrow^{0} C) \\ (\Gamma \,;\, \Delta \longrightarrow^{1} C) \leq (\Gamma' \,;\, \Delta' \longrightarrow^{w} C)\end{array}\right\} \text{ where } \Gamma \subseteq \Gamma' \text{ and } \Delta \subseteq \Delta'.$$

Note that strong sequents never subsume weak sequents.

Obviously, if $s_1 \leq s_2$ and $s_2 \prec s$, then $s_1 \prec s$. It is easy to see that weak sequents model affine logic: this is familiar from embeddings into linear logic that translate affine implications $A \to B$ as $A \multimap (B \otimes \top)$. The collection of inference rules for the forward calculus is in fig. 1. The rules must be read while keeping in mind that they are restricted to subformulas of the goal; precisely, a rule is applied only when the principal formula is a proper decorated subformula of the goal sequent.

The trickiest aspect of these rules are the side conditions (given in parentheses) and the weakness annotations. In order to understand these, it may be useful to think in term of the following property, which we maintain for all rules in order to avoid redundant inferences.

Definition 4. *A rule with conclusion s and premisses $s_1, \ldots, s_n$ is said to satisfy the* irredundancy property *if for no $i \in \{1, \ldots, n\}$, $s_i \leq s$.*

In other words, a rule is irredundant if none of its premisses subsumes the conclusion. Note that this is a local property; we do not discuss here more global redundancy criteria.

The first immediate observation is that binary rules simply take the union of the unrestricted zone from the premisses. The action of the rules on the linear zone is also prescribed by linearity when the sequents are strong ($w = 0$).

The binary additive rule ($\&R$) is applicable in the forward direction when both premisses are weak ($w = 1$), regardless of their linear zone. This is because in this case the linear zones can always be weakened to make them equal. We therefore compute the upper bound ($\sqcup$) of the two multisets: if A occurs n times in Δ and m times in Δ', then it occurs $\max(n, m)$ times in $\Delta \sqcup \Delta'$.

If only one premiss of the binary additive rule is weak, the linear zone of the weak premiss must be included in the linear zone of the other strong premiss. If both premisses are strong, their linear zones must be equal. We abstract the four possibilities in the form of an additive compatibility test.

Definition 5 (additive compatibility). *Given two forward sequents $\Gamma ; \Delta \longrightarrow^{w} C$ and $\Gamma' ; \Delta' \longrightarrow^{w'} C$, their additive zones Δ and Δ' are* additively compatible *given their respective weak-flags, which we write as $\Delta / w \approx \Delta' / w'$, if the following hold:*

$$\begin{array}{llll} \Delta/0 \approx \Delta'/0 & \text{if } \Delta = \Delta' & \Delta/0 \approx \Delta'/1 & \text{if } \Delta' \subseteq \Delta \\ \Delta/1 \approx \Delta'/1 & \text{always} & \Delta/1 \approx \Delta'/0 & \text{if } \Delta \subseteq \Delta' \end{array}$$

For binary multiplicative rules like $\otimes R$, the conclusion is weak if either of the premisses is weak; thus, the weak-flag of the conclusion is a Boolean-or of those of the premisses. Dually, for binary additive rules, the conclusion is weak if both premisses are weak, so we use a Boolean-and to conjoin the weak flags. Most unary rules are oblivious to the weakening decoration, which simply survives from the premiss to the conclusion. The exception is $!R$, for which it is unsound to have a weak conclusion; there is no derivation of $\cdot ; \top \Longrightarrow\, !\top$, for example.

Left rules with weak premisses require some attention. It is tempting to write the "weak" $\otimes L$ rules as:

$$\frac{\Gamma ; \Delta, A \longrightarrow^{1} C}{\Gamma ; \Delta, A \otimes B \longrightarrow^{1} C}\ \otimes L_1 \qquad \frac{\Gamma ; \Delta, B \longrightarrow^{1} C}{\Gamma ; \Delta, A \otimes B \longrightarrow^{1} C}\ \otimes L_2.$$

(Note that the irredundancy property requires that at least one of the operands of $\otimes$ be present in the premiss.) This pair of rules, however, would allow redundant inferences such as:

$$\frac{\Gamma\,;\Delta, A, B \longrightarrow^{1} C}{\Gamma\,;\Delta, A, A \otimes B \longrightarrow^{1} C}\ \otimes L_2.$$

We might as well have consumed both A and B to form the conclusion, and obtained a stronger result. The sensible strategy is: when A and B are both present, they must *both* be consumed. Otherwise, only apply the rule when one operand is present in a weak sequent. A similar observation can be made about all such rules: there is one weakness-agnostic form, and some possible refined forms to account for weak sequents.

Property 6 (irredundancy). *All forward rules satisfy the irredundancy property.* □

The soundness and completeness theorems are both proven by structural induction; we omit the simple proofs. Note that the completeness theorem shows that the forward calculus infers a possibly stronger form of the goal sequent.

Theorem 7 (soundness). *If* $\Gamma\,;\Delta \longrightarrow^{w} C$ *is derivable, then it is sound.*

Theorem 8 (completeness). *If* $\Gamma\,;\Delta \Longrightarrow C$ *is derivable, then there exists a derivable forward sequent* $\Gamma'\,;\Delta' \longrightarrow^{w} C$ *such that* $\left(\Gamma'\,;\Delta' \longrightarrow^{w} C\right) \prec \left(\Gamma\,;\Delta \Longrightarrow C\right)$.

4 Focused derivations

Search using the backward calculus can always apply invertible rules eagerly in any order as there always exists a proof that goes through the premisses of the invertible rule. Andreoli pointed out [1] that a similar and dual feature exists for non-invertible rules also: it is enough for completeness to apply a sequence of non-invertible rules eagerly in one atomic operation, as long as the corresponding connectives are of the same *synchronous* nature.

In classical linear logic the synchronous or asynchronous nature of a given connective is identical to its polarity; the negative connectives (&, $\top$, ⅋, $\bot$, $\forall$) are asynchronous, and the positive connectives ($\otimes$, $\mathbf{1}$, $\oplus$, $\mathbf{0}$, $\exists$) are synchronous. The nature of intuitionistic connectives, though, must be derived without an appeal to polarity, which is alien to the constructive and judgmental philosophy underlying the logic. We derive this nature by examining the rules and phases of search: an asynchronous connective is one for which decomposition is complete in the active phase; a synchronous connective is one for decomposition is complete in the focused phase. This definition happens to coincide with polarities for classical linear logic. Our reconstruction of focusing for intuitionistic linear logic can be seen as a refinement of Howe's approach [3]: our calculus differs in having a precise definition the classes of connectives and no overlap in the inference rules. As a result, our completeness proof is considerably simpler, being a direct consequence of cut-elimination.

As our backward linear sequent calculus is two-sided, we have left- and right-synchronous and asynchronous connectives. For non-atomic propositions a left-synchronous connective is right-asynchronous, and a left-asynchronous connective right-synchronous; this appears to be universal in well-behaved logics. We define the notations in the table below

symbol	meaning
P	left-synchronous (&, $\top$, $\multimap$, p)
Q	right-synchronous ($\otimes$, $\mathbf{1}$, !, p)
L	left-asynchronous ($\otimes$, $\mathbf{1}$, !)
R	right-asynchronous (&, $\top$, $\multimap$)

The backward focusing calculus consists of three kinds of sequents; *right-focal sequents* of the form $\Gamma\,;\Delta \gg A$ (A under focus), *left-focal sequents* of the form $\Gamma\,;\Delta\,;A \ll Q$, and *active sequents* of the form $\Gamma\,;\Delta\,;\Omega \Longrightarrow C$. Γ indicates the unrestricted zone as usual, Δ contains *only* left-synchronous propositions, and Ω is an ordered sequence of propositions (of arbitrary nature).

The active phase is entirely deterministic: it operates on the right side of the active sequent until it becomes right-synchronous, i.e., of the form $\Gamma\,;\Delta\,;\Omega \Longrightarrow Q$. Then the propositions in Ω are decomposed in order from right to left. The order of Ω is used solely to avoid spurious non-deterministic choices. Eventually the sequent is reduced to the form $\Gamma\,;\Delta\,;\cdot \Longrightarrow Q$, which we call *neutral sequents*.

A focusing phase is launched from a neutral sequent by selecting a proposition from Γ, Δ or the right hand side. This focused formula is decomposed until the top-level connective becomes asynchronous. Then we enter an active phase for the previously focused proposition.

Atomic propositions and modal operators need a special mention. Andreoli observed in [1] that it is sufficient to assign arbitrarily a synchronous or asynchronous nature to the atoms as long as duality is preserved; here, the asymmetric nature of the intuitionistic sequents suggests that they should be synchronous, as explained below. The modal connectives were treated by Howe as neither synchronous nor asynchronous, If the left-focal formula is an atom, then the sequent is initial iff the linear zone Δ is empty *and* the right hand side matches the focused formula; this gives the focused version of the "init" rule. If an atom has right-focus, however, it is not enough to simply check that the left matches the right, as there might be some pending decompositions; consider eg. $\cdot\,;p \mathbin{\&} q \gg q$. Focus is therefore blurred in this case, and we correspondingly disallow a right atom in a neutral sequent from gaining focus.

The other subtlety is with the $!R$ rule: although ! is right synchronous, the $!R$ rule cannot maintain focus on the operand. If this were forced, there could be no focused proof of $!(A \otimes B) \multimap !(B \otimes A)$, for example. This is because there is a hidden transition from the truth of $!\,A$ to the validity of A which in turn reduces to the truth of A (see [6]). The first is synchronous, the second asynchronous, so the exponential has aspects of both. Girard has made a similar observation that exponentials are composed of one micro-connective to change polarity, and another to model a given behavior [14, Page 114]; this observation extends to other modal operators, such as why-not (?) of JILL [6] or the lax modality of CLF [7].

The full set of rules is in fig. 2. Soundness of this calculus is rather an obvious property— forget the distinction between Δ and Ω, elide the focus and blur rules, and the original backward calculus appears. For completeness of the focusing calculus, we proceed by interpreting every backward sequent as an active sequent in the focusing calculus, then showing that the backward rules are admissible in the focusing calculus. This proof relies on admissibility of cut in the focusing calculus. Because a non-

right-focal

$$\frac{\Gamma ; \Delta_1 \gg A \quad \Gamma ; \Delta_2 \gg B}{\Gamma ; \Delta_1, \Delta_2 \gg A \otimes B}\ \otimes R$$

$$\frac{}{\Gamma ; \cdot \gg \mathbf{1}}\ \mathbf{1}R \qquad \frac{\Gamma ; \cdot ; \cdot \Longrightarrow A}{\Gamma ; \cdot \gg\, !A}\ !R$$

left-focal

$$\frac{}{\Gamma ; \cdot ; p \ll p}\ \text{init} \qquad \frac{\Gamma ; \Delta ; A_i \ll Q}{\Gamma ; \Delta ; A_1 \,\&\, A_2 \ll Q}\ \& L_i$$

$$\frac{\Gamma ; \Delta_1 ; B \ll Q \quad \Gamma ; \Delta_2 \gg A}{\Gamma ; \Delta_1, \Delta_2 ; A \multimap B \ll Q}\ \multimap R$$

focus

$$\frac{\Gamma ; \Delta ; P \ll Q}{\Gamma ; \Delta, P ; \cdot \Longrightarrow Q}\ \text{lf}$$

$$\frac{\Gamma, A ; \Delta ; A \ll Q}{\Gamma, A ; \Delta ; \cdot \Longrightarrow Q}\ \text{copy}$$

$$\frac{\Gamma ; \Delta \gg Q \quad Q \text{ non-atomic}}{\Gamma ; \Delta ; \cdot \Longrightarrow Q}\ \text{rf}$$

right-active

$$\frac{\Gamma ; \Delta ; \Omega \Longrightarrow A \quad \Gamma ; \Delta ; \Omega \Longrightarrow B}{\Gamma ; \Delta ; \Omega \Longrightarrow A \,\&\, B}\ \& R$$

$$\frac{}{\Gamma ; \Delta ; \Omega \Longrightarrow \top}\ \top R \qquad \frac{\Gamma ; \Delta ; \Omega \cdot A \Longrightarrow B}{\Gamma ; \Delta ; \Omega \Longrightarrow A \multimap B}\ \multimap R$$

left-active

$$\frac{\Gamma ; \Delta ; \Omega \cdot A \cdot B \Longrightarrow Q}{\Gamma ; \Delta ; \Omega \cdot A \otimes B \Longrightarrow Q}\ \otimes L$$

$$\frac{\Gamma ; \Delta ; \Omega \Longrightarrow Q}{\Gamma ; \Delta ; \Omega \cdot \mathbf{1} \Longrightarrow Q}\ \mathbf{1}L \qquad \frac{\Gamma, A ; \Delta ; \Omega \Longrightarrow Q}{\Gamma ; \Delta ; \Omega \cdot\, !A \Longrightarrow Q}\ !L$$

$$\frac{\Gamma ; \Delta, P ; \Omega \Longrightarrow Q}{\Gamma ; \Delta ; \Omega \cdot P \Longrightarrow Q}\ \text{act}$$

blur

$$\frac{\Gamma ; \Delta ; L \Longrightarrow Q}{\Gamma ; \Delta ; L \ll Q}\ \text{lb}$$

$$\frac{\Gamma ; \Delta ; \cdot \Longrightarrow R}{\Gamma ; \Delta \gg R}\ \text{rb} \qquad \frac{\Gamma ; \Delta ; \cdot \Longrightarrow p}{\Gamma ; \Delta \gg p}\ \text{rb}^*$$

Fig. 2: Backward linear focusing calculus

atomic left-synchronous proposition is right-asynchronous, a left-focal sequent needs to match only an active sequent in a cut; similarly for right-synchronous propositions. Active sequents should match other active sequents, however. Cuts destroy focus, as they generally require commutations spanning phase boundaries; the products of a cut are therefore active.

The proof needs two key lemmas: the first notes that permuting the ordered context doesn't affect provability, as the ordered context does not mirror any deep non-commutativity in the logic. This lemma thus allows cutting formulas from anywhere inside the ordered context, and also to re-order the context when needed.

Lemma 9. *If $\Gamma ; \Delta ; \Omega \Longrightarrow C$, then $\Gamma ; \Delta ; \Omega' \Longrightarrow C$ for any permutation Ω' of Ω.* □

The other lemma shows that left-active rules can be applied even if the right-hand side is not synchronous. This lemma is vital for commutative cuts.

Lemma 10. *The following variants of the left-active rules are admissible*

$$\frac{\Gamma ; \Delta, P ; \Omega \Longrightarrow C}{\Gamma ; \Delta ; \Omega \cdot P \Longrightarrow C} \qquad \frac{\Gamma ; \Delta ; \Omega \cdot A \cdot B \Longrightarrow C}{\Gamma ; \Delta ; \Omega \cdot A \otimes B \Longrightarrow C} \qquad \frac{\Gamma ; \Delta ; \Omega \Longrightarrow C}{\Gamma ; \Delta ; \Omega \cdot \mathbf{1} \Longrightarrow C} \qquad \frac{\Gamma, A ; \Delta ; \Omega \Longrightarrow C}{\Gamma ; \Delta ; \Omega \cdot\, !A \Longrightarrow C}$$

Theorem 11 (cut). *If*

1. *$\Gamma ; \Delta \gg A$ and:*
   (a) *$\Gamma ; \Delta' ; \Omega \cdot A \Longrightarrow C$ then $\Gamma ; \Delta, \Delta' ; \Omega \Longrightarrow C$.*

(b) $\Gamma\,;\Delta',A\,;\Omega \Longrightarrow C$ *then* $\Gamma\,;\Delta,\Delta'\,;\Omega \Longrightarrow C$.

2. $\Gamma\,;\cdot \gg A$ *and* $\Gamma,A\,;\Delta\,;\Omega \Longrightarrow C$ *then* $\Gamma\,;\Delta\,;\Omega \Longrightarrow C$.
3. $\Gamma\,;\Delta\,;\Omega \Longrightarrow A$ *and:*
   (a) $\Gamma\,;\Delta'\,;A \ll Q$ *then* $\Gamma\,;\Delta,\Delta'\,;\Omega \Longrightarrow Q$.
   (b) $\Gamma\,;\Delta'\,;\Omega'\cdot A \Longrightarrow C$ *then* $\Gamma\,;\Delta,\Delta'\,;\Omega\cdot\Omega' \Longrightarrow C$.
   (c) $\Gamma\,;\Delta',A\,;\Omega' \Longrightarrow C$ *then* $\Gamma\,;\Delta,\Delta'\,;\Omega\cdot\Omega' \Longrightarrow C$.
4. $\Gamma\,;\cdot\,;\cdot \Longrightarrow A$ *and:*
   (a) $\Gamma,A\,;\Delta\,;\Omega \Longrightarrow C$ *then* $\Gamma\,;\Delta\,;\Omega \Longrightarrow C$.
   (b) $\Gamma,A\,;\Delta \gg B$ *then* $\Gamma\,;\Delta \gg B$.
5. $\Gamma\,;\Delta\,;B \ll A$ *and:*
   (a) $\Gamma\,;\Delta'\,;A \Longrightarrow Q$ *then* $\Gamma\,;\Delta,\Delta'\,;B \ll Q$.
   (b) $\Gamma\,;\Delta',A\,;\cdot \Longrightarrow Q$ *then* $\Gamma\,;\Delta,\Delta'\,;B \ll Q$.
6. $\Gamma\,;\Delta\,;\cdot \Longrightarrow A$ *and* $\Gamma\,;\Delta',A \gg B$ *then* $\Gamma\,;\Delta,\Delta' \gg B$.

Proof (sketch). By lexicographic induction on the given derivations. The argument is lengthy rather than complex, and is an adaptation of similar structural cut-admissibility proofs in eg. [6]. □

Theorem 12 (completeness).
If $\Gamma\,;\Delta \Longrightarrow C$ *and* Ω *is any serialization of* Δ, *then* $\Gamma\,;\cdot\,;\Omega \Longrightarrow C$.

Proof (sketch). First show that all ordinary rules are admissible in the focusing system using cut. Proceed by induction on the derivation of $\mathcal{D} :: \Gamma\,;\Delta \Longrightarrow C$, splitting cases on the last applied rule, using cut and lem. 9 as required. A representative case (for $\otimes R$):

$$\mathcal{D} = \frac{\mathcal{D}_1 :: \Gamma\,;\Delta \Longrightarrow A \quad \mathcal{D}_2 :: \Gamma\,;\Delta' \Longrightarrow B}{\Gamma\,;\Delta,\Delta' \Longrightarrow A\otimes B}\ \otimes R$$

Let Ω and Ω' be serializations of Δ and Δ' respectively; by the induction hypothesis on $\mathcal{D}_1$ and $\mathcal{D}_2$, we have $\Gamma\,;\cdot\,;\Omega \Longrightarrow A$ and $\Gamma\,;\cdot\,;\Omega' \Longrightarrow B$. Now, it is easy to show that $\Gamma\,;\cdot\,;A\cdot B \Longrightarrow A\otimes B$. The result follows by the use of cut twice, for A and B in the active context respectively, to get $\Gamma\,;\cdot\,;\Omega\cdot\Omega' \Longrightarrow A\otimes B$, and then noting that any serialization of Δ,Δ' is a permutation of $\Omega\cdot\Omega'$. □

5 Forward focusing

We now construct the forward version of the focusing calculus. Intermediate sequents in the eager active and focusing phases must not be added to the database of derived sequents; instead, we should store just the neutral sequents—that is, of the form $\Gamma\,;\Delta\,;\cdot \Longrightarrow Q$—at the phase boundaries. We therefore first construct derived rules for neutral sequents that make the intermediate focal and active sequents irrelevant.

For any given synchronous subformula, the derived inferences for that subformula correspond to a single pair of focal and active phases; Andreoli called them *bipoles* [2]. However, there are important differences between backward reasoning bipoles and their forward analogue: as shown in thm. 8, the forward calculus generates stronger forms of sequents than in the corresponding backward proof. Therefore, not every branch of the backward bipole will be available in the forward direction. The forward derived rules therefore need some additional mechanism in the internal nodes to handle these cases.

We still adapt the essential idea of bipoles of viewing every proposition as a relation between the conclusion of the bipole and its possible premisses at the leaves of the bipole. This relational interpretation gives us the derived rules corresponding to the proposition; the premisses and conclusions of these derived rules are neutral sequents, which we indicate by means of a double-headed sequent arrow ($\longrightarrow\!\!\!\!\!\rightarrow$).

Each relation R takes as input the premisses of the bipole, $s_1 \cdot s_2 \cdots s_n$ (written Σ), and constructs the relevant portion of a conclusion sequent s; we write this as $R[\Sigma] \hookrightarrow s$. There are three classes of these relations:

1. Right focal relations for the focus formula A, written $\mathsf{foc}^+_{\Downarrow}(A)$.
2. Left focal relations for the focus formula A, written $\mathsf{foc}^-_{\Downarrow}(A)$.
3. Active relations, written $\mathsf{act}_{\Downarrow}(\Gamma\,;\Delta\,;\Omega \Longrightarrow \gamma)$, where γ is either $\cdot$ or C.

The focal relations are understood as defining derived rules corresponding to a given proposition. The conclusion of these derived rules are themselves neutral sequents. For a right focal relation $\mathsf{foc}^+_{\Downarrow}(Q)$, the corresponding derived rule is:

$$\frac{\Sigma \quad \left(\mathsf{foc}^+_{\Downarrow}(Q)[\Sigma] \hookrightarrow \Gamma\,;\Delta \longrightarrow\!\!\!\!\!\rightarrow^{w} \cdot\right)}{\Gamma\,;\Delta \longrightarrow\!\!\!\!\!\rightarrow^{w} Q}\ \mathsf{foc}^+_{\Downarrow}$$

Similarly, for negative propositions, we have two rules, depending on whether the focused proposition is a heavy subformula of the goal sequent or not.

$$\frac{\Sigma \quad \left(\mathsf{foc}^-_{\Downarrow}(P)[\Sigma] \hookrightarrow \Gamma\,;\Delta \longrightarrow\!\!\!\!\!\rightarrow^{w} Q\right)}{\Gamma\,;\Delta, P \longrightarrow\!\!\!\!\!\rightarrow^{w} Q}\ \mathsf{foc}^-_{\Downarrow} \qquad \frac{\Sigma \quad \left(\mathsf{foc}^-_{\Downarrow}(A)[\Sigma] \hookrightarrow \Gamma\,;\Delta \longrightarrow\!\!\!\!\!\rightarrow^{w} Q\right)}{\Gamma \cup \{A\}\,;\Delta \longrightarrow\!\!\!\!\!\rightarrow^{w} Q}\ !\,\mathsf{foc}^-_{\Downarrow}$$

As before, these derived rules are understood to contain only signed subformulas of the goal sequent. The active relations essentially replay the active rules of the backward focusing calculus, except they also account for weak sequents as needed.

For lack of space we leave out the details of the definition of these relations; they can be found in the accompanying technical report [15]. Instead, we shall give an example. Consider the negative principal subformula $P = p \,\&\, q \multimap r \,\&\, (s \otimes t)$ and the three input sequents $\Gamma_1\,;\Delta_1 \longrightarrow\!\!\!\!\!\rightarrow^1 p$, $\Gamma_2\,;\Delta_2 \longrightarrow\!\!\!\!\!\rightarrow^0 q$, and $\Gamma_3\,;\Delta_3, s \longrightarrow\!\!\!\!\!\rightarrow^1 Q$, named s_1, s_2, and s_3 respectively. By the definition of $\mathsf{foc}^-_{\Downarrow}$:

$$\mathsf{foc}^-_{\Downarrow}(P)[s_3 \cdot s_1 \cdot s_2] \hookrightarrow \Gamma_3 \cup \Gamma_1 \cup \Gamma_2\,;\Delta_3, \Delta_2 \longrightarrow\!\!\!\!\!\rightarrow^1 Q \qquad \text{if } t \notin \Delta_3 \text{ and } \Delta_1 \subseteq \Delta_2$$

In other words, the instance of the full derived rule for P matched against the given sequents stands for the following derived rule of inference specialized to this scenario:

$$\frac{\Gamma_1\,;\Delta_1 \longrightarrow\!\!\!\!\!\rightarrow^1 p \quad \Gamma_2\,;\Delta_2 \longrightarrow\!\!\!\!\!\rightarrow^0 q \quad \Gamma_3\,;\Delta_3, s \longrightarrow\!\!\!\!\!\rightarrow^1 Q \quad (t \notin \Delta_3) \quad (\Delta_1 \subseteq \Delta_2)}{\Gamma_3 \cup \Gamma_1 \cup \Gamma_2\,;\Delta_3, \Delta_2, P \longrightarrow\!\!\!\!\!\rightarrow^1 Q}$$

The proofs of soundness and completeness of the forward focusing calculus with respect to the backward focusing calculus are are in [15]. Soundness is shown by simple structural induction on the $\mathsf{foc}^+_{\Downarrow}$, $\mathsf{foc}^-_{\Downarrow}$ and $\mathsf{act}_{\Downarrow}$ derivations. Completeness is a rather more complex result because the forward and backward focused proofs are not in bijection. The essential idea of the proof is to define a complete calculus of backward derived rules, and prove the calculus of forward derived rules complete with respect to this intermediate calculus.

Theorem 13 (soundness). *If $\Gamma\,;\Delta \longrightarrow\!\!\!\!\rightarrow^{w} Q$ is derivable, then it is sound.* □

Theorem 14 (completeness). *If $\Gamma\,;\Delta\,;\cdot \Longrightarrow Q$ is derivable, then there exists a derivable focused sequent $\Gamma'\,;\Delta' \longrightarrow\!\!\!\!\rightarrow^{w} Q$ such that $\left(\Gamma'\,;\Delta' \longrightarrow\!\!\!\!\rightarrow^{w} Q\right) \prec \left(\Gamma\,;\Delta \Longrightarrow Q\right)$.* □

5.1 The focused inverse method

What remains is to implement the inverse method search strategy that uses the forward focusing calculus. We only briefly sketch the method here, as the implementation issues are out of the scope of this paper, and have been detailed in a related paper [9]. The inverse method consists of three essential components – the *database* of generated sequents, the library of *rules* that can be applied to sequents to generate new sequents, and the main loop or *engine*. Rules are constructed by naming all subformulas of the goal sequent with fresh propositional labels, and specializing the inference rules of the full logic to principal uses of the subformula labels; the general rules are then discarded. This procedure is key to giving the inverse method a goal direction, as the search space is constrained to subformulas of the goal. Traditionally the library of rules is considered static during a given search, but as we describe in [9], it is beneficial, especially in the first-order extension, to allow the library of rules to be extended during search with partial applications– a form of memoization. The inputs for these rules are drawn from the database of computed sequents. At the start of search, this database contains just the initial sequents, which are determined by considering all *atomic* subformulas that are both positively and negatively occurring in the goal sequent. The engine repeatedly selects sequents from the database, and applies rules from the library to generate new sequents; if these new sequents are not subsumed by any sequent derived earlier, they are inserted in to the database. Completeness of the search strategy is guaranteed by using a fair selection (i.e., equivalent to breadth-first search) of sequents from the database in order to generate new sequents.

The primary issue in the presence of focusing is what propositions to generate rules for. As the calculus of derived rules has only neutral sequents as premisses and conclusions, we need only generate rules for propositions that occur in neutral sequents; we call them *frontier propositions*. To find the frontier propositions in a goal sequent, we simply abstractly replay the focusing and active phases to identify the phase transitions. Each transition from an active to a focal phase produces a frontier proposition. Formally, we define two generating functions, f (focal) and a (active), from signed propositions to multisets of frontier propositions. None of the logical constants are in the frontier, for the conclusions of rules such as $\top R$ and $\mathbf{1}R$ are easy to predict, and can be generated as needed. Similarly we do not count a negative focused atomic proposition in the frontier as we know that the conclusion of the init rule needs to have the form $\Gamma\,;\cdot\,;p \ll p$; this restricts the collection of spurious initial sequents that are not possible in a focused proof. The steps in the calculation are shown in figure 3; as a simple example, $f(p \,\&\, q \multimap r \,\&\, (s \otimes t))^{-} = p, q, s, t$.

$$f(p)^- = \emptyset \quad f(p)^+ = a(p)^\pm = \{p\} \quad f(\mathbf{1})^\pm = a(\mathbf{1})^\pm = \emptyset \quad f(\top)^\pm = a(\top)^\pm = \emptyset$$

$$\begin{array}{ll} f(A \otimes B)^- = a(A \otimes B)^- & f(A \otimes B)^+ = f(A)^+, f(B)^+ \\ a(A \otimes B)^- = a(A)^-, a(B)^- & a(A \otimes B)^+ = f(A \otimes B)^+, A \otimes B \end{array}$$

$$\begin{array}{ll} f(A \mathbin{\&} B)^- = f(A)^-, f(B)^- & f(A \mathbin{\&} B)^+ = a(A \mathbin{\&} B)^+ \\ a(A \mathbin{\&} B)^- = f(A \mathbin{\&} B)^-, A \mathbin{\&} B & a(A \mathbin{\&} B)^+ = a(A)^+, a(B)^+ \end{array}$$

$$\begin{array}{ll} f(A \multimap B)^- = f(A)^+, f(B)^- & f(A \multimap B)^+ = a(A \multimap B)^+ \\ a(A \multimap B)^- = f(A \multimap B)^-, A \multimap B & a(A \multimap B)^+ = a(A)^-, a(B)^+ \end{array}$$

$$f(!\,A)^- = a(!\,A)^- \quad f(!\,A)^+ = a(A)^+ \quad a(!\,A)^- = a(A)^- \quad a(!\,A)^+ = f(A)^+, !\,A$$

Fig. 3: calculating frontier propositions

Definition 15 (frontier). *Given a goal $\Gamma \,;\Delta \,; \cdot \Longrightarrow Q$ (which is neutral), its frontier contains:*

i. all (top-level) propositions in Γ, Δ, Q;
ii. for any $A \in \Gamma, \Delta$, the collection $f(A)^-$; and
iii. the collection $f(Q)^+$.

Property 16 (neutral subformula property). *In any backward focused proof, all neutral sequents consist only of frontier propositions of the goal sequent.* □

In the preparatory phase for the inverse method, we calculate the frontier propositions of the goal sequent. There is no need to generate initial sequents separately, as the executions of negative atoms in the frontier directly give us the necessary initial sequents. The general design of the main loop of the prover and the argument for its completeness are fairly standard [8, 10]; we use a lazy refinement of this basic design [9] that is ideal for multi-premiss rules.

5.2 Some experimental results

We have implemented an expanded version of the forward focusing calculus as a certifying[1] inverse method prover for intuitionistic linear logic, including the missing connectives $\oplus$, $\mathbf{0}$, and the lax modality.[2] Table 1 contains a running-time comparison of the focusing prover (**F**) against a non-focusing version (**NF**) of the prover (directly implementing the calculus of sec. 3), and Tammet's Gandalf "nonclassical" distribution that includes a pair of (non-certifying) provers for classical linear logic, one (**Gr**) using a refinement of Mints' resolution system for classical linear logic [11, 10], and the other (**Gt**) using a backward Tableaux-based strategy. Neither of these provers incorporates focusing. The test problems ranged from simple stateful encodings such as blocksworld or change-machines, to more complex problems such as encoding of affine logic

[1] By *certifying*, we mean that it produces independently verifiable proof objects.
[2] Available from the first author's web page at `http://www.cs.cmu.edu/~kaustuv/`

Test	NF	F	Gt	Gr
`blocks-world`	0.02 s	≤ 0.01 s	13.51 s	0.03 s
`change`	3.20 s	≤ 0.01 s	—	0.63 s
`affine1`	0.01 s	≤ 0.01 s	0.03 s	≤ 0.01 s
`affine2`	≈ 12 m	1.21 s	—	—
`qbf1`	0.03 s	≤ 0.01 s	—	2.40 s
`qbf2`	0.04 s	≤ 0.01 s	—	42.34 s
`qbf3`	≈ 35 m	0.53 s	—	—

All measurements are wall-clock times on an unloaded computer with a 2.80GHz Pentium 4 processor, 512KB L1 cache and 1GB of main memory; "—" denotes unsuccessful proof within ≈ ten hours.

Table 1: some experimental results.

problems, and translations of various quantified Boolean formulas using the algorithm in [16]. Focusing was faster in every case, with an average speedup of about three orders of magnitude over the non-focusing version.

6 Embedding non-linear logics

6.1 Intuitionistic logic

When we move from intuitionistic to intuitionistic linear logic, we gain a lot of expressive power. Nonetheless, many problems, even if posed in linear logic, have significant non-linear components or subproblems. Standard translations into linear logic, however, have the problem that any focusing properties enjoyed by the source are lost in the translation. In a focusing system for intuitionistic logic, as hinted to by Howe [3] and briefly considered below, a quite deterministic proof with, say, one phase of focusing, will be decomposed into many small phases, leading to a large loss in efficiency. Fortunately, it is possible to translate intuitionistic logic in a way that preserves focusing. To illustrate, consider a minimal intuitionistic propositional logic with connectives $\{\wedge, \mathtt{t}, \supset\}$. The focusing system for this logic has three kinds of sequents, $\Gamma \gg_I A$ (right-focal), $\Gamma \,;\, A \ll_I Q$ (left-focal), and $\Gamma \,;\, \Omega \Longrightarrow_I C$ (active), with $\supset$ treated as right-synchronous, and $\wedge$ as *both* (right-) synchronous and asynchronous. The meta-variables P, Q, L and R are used in the spirit of section 4; that is, P for left-synchronous $\{\wedge, \mathtt{t}, \supset, p\}$, Q for right-synchronous $\{\wedge, \mathtt{t}, p\}$, L for left-asynchronous $\{\wedge, \mathtt{t}\}$, and R for right-asynchronous $\{\wedge, \mathtt{t}, \supset\}$. Q^* means that Q is not atomic, i.e., just containing $\{\wedge, \mathtt{t}\}$.

$$\frac{}{\Gamma \,;\, p \ll_I p} \qquad \frac{\Gamma \,;\, A_i \ll_I Q}{\Gamma \,;\, A_1 \wedge A_2 \ll_I Q} \qquad \frac{\Gamma \,;\, B \ll_I Q \quad \Gamma \gg_I A}{\Gamma \,;\, A \supset B \ll_I Q} \qquad \frac{\Gamma \gg_I A \quad \Gamma \gg_I B}{\Gamma \gg_I A \wedge B} \qquad \frac{}{\Gamma \gg_I \mathtt{t}}$$

$$\frac{\Gamma \,;\, \Omega \cdot A \cdot B \Longrightarrow_I Q}{\Gamma \,;\, \Omega \cdot A \wedge B \Longrightarrow_I Q} \qquad \frac{\Gamma \,;\, \Omega \Longrightarrow_I A \quad \Gamma \,;\, \Omega \Longrightarrow_I B}{\Gamma \,;\, \Omega \Longrightarrow_I A \wedge B} \qquad \frac{}{\Gamma \,;\, \Omega \Longrightarrow_I \mathtt{t}} \qquad \frac{\Gamma \,;\, \Omega \cdot A \Longrightarrow_I B}{\Gamma \,;\, \Omega \Longrightarrow_I A \supset B}$$

$$\frac{\Gamma, P \,;\, \Omega \Longrightarrow_I Q}{\Gamma \,;\, \Omega \cdot P \Longrightarrow_I Q}\,\text{act} \qquad \frac{\Gamma \gg_I Q^*}{\Gamma \,;\, \cdot \Longrightarrow_I Q^*} \qquad \frac{\Gamma \,;\, P \ll_I Q}{\Gamma, P \,;\, \cdot \Longrightarrow_I Q} \qquad \frac{\Gamma \,;\, \cdot \Longrightarrow_I R}{\Gamma \gg_I R} \qquad \frac{\Gamma \,;\, L \Longrightarrow_I Q}{\Gamma \,;\, L \ll_I Q}$$

The translation is modal with two phases: A (active) and F (focal). A positive focal $\wedge$ is translated as $\otimes$, and the duals as &. For every use of the act rule, the corresponding

$$F(p)^- = p \quad F(p)^+ = p \quad A(p)^- = !\,p \quad A(p)^+ = p$$

$$\begin{array}{ll} F(A \wedge B)^- = F(A)^- \mathbin{\&} F(B)^- & F(A \wedge B)^+ = F(A)^+ \otimes F(B)^+ \\ A(A \wedge B)^- = A(A)^- \otimes A(B)^- & A(A \wedge B)^+ = A(A)^+ \mathbin{\&} A(B)^+ \end{array}$$

$$F(\mathtt{t})^- = \top \quad F(\mathtt{t})^+ = \mathbf{1} \quad A(\mathtt{t})^- = \mathbf{1} \quad A(\mathtt{t})^+ = \top$$

$$\begin{array}{ll} F(A \supset B)^- = F(A)^+ \multimap F(B)^- & F(A \supset B)^+ = A(A \supset B)^+ \\ A(A \supset B)^- = !\,F(A \supset B)^- & A(A \supset B)^+ = A(A)^- \multimap A(B)^+ \end{array}$$

Fig. 4: embedding intuitionistic logic

translation phase affixes an exponential; the phase-transitions in the image of the translation exactly mirror those in the source. The details of the translation are in figure 4. It is easily shown that these translations preserve the focusing structure of proofs.

Property 17 (preservation of the structure of proofs).

1. *If $\Gamma \gg_I A$, then $F(\Gamma)^- \,;\, \cdot \gg F(A)^+$.*
2. *If $\Gamma \,;\, A \ll_I Q$, then $F(\Gamma)^- \,;\, \cdot \,;\, F(A)^- \ll F(Q)^+$.*
3. *If $\Gamma \,;\, \Omega \Longrightarrow_I Q$, then $F(\Gamma)^- \,;\, \cdot \,;\, A(\Omega)^- \Longrightarrow F(Q)^+$.*
4. *If $\Gamma \,;\, \Omega \Longrightarrow_I R$, then $F(\Gamma)^- \,;\, \cdot \,;\, A(\Omega)^- \Longrightarrow A(R)^+$.* □

The reverse translation, written $-^o$, is trivial: simply erase all !s, rewrite & and ⊗ as ∧, ⊤ and **1** as t, and ⊸ as ⇒.

Property 18 (soundness).

1. *If $\Gamma \,;\, \cdot \gg A$, then $\Gamma^o \gg_I A^o$.*
2. *If $\Gamma \,;\, \cdot \,;\, A \ll Q$, then $\Gamma^o \,;\, A^o \ll_I Q^o$.*
3. *If $\Gamma \,;\, \cdot \,;\, \Omega \Longrightarrow C$, then $\Gamma^o \,;\, \Omega^o \Longrightarrow_I C^o$.* □

An important feature of this translation is that only (certain) negative atoms and implications are !-affixed; this is related to a similar observation by Dyckhoff that the ordinary propositional intuitionistic logic has a contraction-free sequent calculus that duplicates only negative atoms and implications [17]. It is also important to note that this translation extends easily to handle the disjunctions ∨ and ⊥ (in the source) and ⊕ and **0** in the target logic; this naturality is not as obvious for Howe's translation [3].

6.2 Classical Horn formulas

A related issue arises with respect to (non-linear) Horn logic. In complex specifications that employ linearity, there are often significant sub-specifications that lie in the Horn fragment. Unfortunately, the straightforward inverse method is quite inefficient on Horn formulas, something already noticed by Tammet [10]. So his prover switches between hyperresolution for Horn and near-Horn formulas and the inverse method for other propositions.

With focusing, this becomes entirely unnecessary. Our focused inverse method for intuitionistic linear logic, when applied to a classical, non-linear Horn formula, will exactly behave as classical hyperresolution. This remarkable property gives further evidence to the power of focusing as a technique for forward theorem proving.

A propositional Horn clause has the form $p_1 \supset \cdots \supset p_n \supset p$ where all p_i and p are atomic. A Horn theory Ψ is just a set of Horn clauses. This can easily be generalized to include conjunction and truth. The results in this section extend also to the first-order case, where Horn formulas allow outermost universal quantification.

$$\frac{p_1 \quad p_2 \quad \cdots \quad p_n}{p}\ \text{hyper}$$

The hyperresolution strategy on this framework is essentially just forward reasoning with rule set "hyper" for any $p_1 \supset \cdots \supset p_n \supset p \in \Psi$. Note that these will be unit clauses if $n = 0$. If we translate every clause $p_1 \supset \cdots \supset p_n \supset p$ as $!(p_1 \multimap \cdots \multimap p_n \multimap p)$, it is easy to see that the derived rules associated with the results of the translation are exactly the hyperresolution rules.

7 Conclusion

We have presented the design of an inverse method theorem prover for propositional intuitionistic linear logic and have demonstrated through experimental results that focusing represents a highly significant improvement. Though elided here, the results persist in the presence of a lax modality [7], and extend to the first-order case as shown by the authors in a related paper [9], which also contains many more details on the implementation and a more thorough empirical evaluation.

Our methods derived from focusing can be applied directly and more easily to classical linear logic and (non-linear) intuitionistic logic, also yielding focused inverse method provers. While we do not have an empirical evaluation of such provers, the reduction in the complexity of the search space is significant. We therefore believe that focusing is a nearly universal improvement to the inverse method and should be applied as a matter of course, possibly excepting only (non-linear) classical logic.

In future work we plan to add higher-order and linear terms in order to obtain a theorem prover for all of CLF [7]. The main obstacles will be to develop feasible algorithms for unification and to integrate higher-order equational constraints. We are also interested in exploring if model-checking techniques could help to characterize the shape of the linear zone that could arise in a backward proof in order to further restrict forward inferences.

Finally, we plan a more detailed analysis of connections with a bottom-up logic programming interpreter for the LO fragment of classical linear logic [18]. This fragment, which is in fact affine, has the property that the unrestricted context remains constant throughout a derivation, and incorporates focusing at least partially via a backchaining rule. It seems plausible that our prover might simulate their interpreter when LO specifications are appropriately translated into intuitionistic linear logic, similar to the translation of classical Horn clauses.

Acknowledgments: We thank Kevin Watkins for illuminating discussions on the topic of focused derivations, and the anonymous referees of this and an earlier version of this paper for their helpful suggestions.

References

1. Andreoli, J.M.: Logic programming with focusing proofs in linear logic. Journal of Logic and Computation **2** (1992) 297–347
2. Andreoli, J.M.: Focussing and proof construction. Annals of Pure and Applied Logic **107** (2001) 131–163
3. Howe, J.M.: Proof Search Issues in Some Non-Classical Logics. PhD thesis, University of St. Andrews (1998)
4. Girard, J.Y.: Linear logic. Theoretical Computer Science **50** (1987) 1–102
5. Barber, A.: Dual Intuitionistic Linear Logic. Technical Report ECS-LFCS-96-347, University of Edinburgh (1996)
6. Chang, B.Y.E., Chaudhuri, K., Pfenning, F.: A judgmental analysis of linear logic. Technical Report CMU-CS-03-131R, Carnegie Mellon University (2003)
7. Cervesato, I., Pfenning, F., Walker, D., Watkins, K.: A concurrent logical framework I & II. Technical Report CMU-CS-02-101 and 102, Department of Computer Science, Carnegie Mellon University (2002) Revised May 2003.
8. Degtyarev, A., Voronkov, A.: The Inverse Method. In: Handbook of Automated Reasoning. MIT Press (2001) 179–272
9. Chaudhuri, K., Pfenning, F.: A focusing inverse method theorem prover for first-order linear logic. In: Proceedings of the 20th International Conference on Automated Deduction (CADE-20). (2005) To appear.
10. Tammet, T.: Resolution, inverse method and the sequent calculus. In: Proceedings of KGC'97, Springer-Verlag LNCS 1289 (1997) 65–83
11. Mints, G.: Resolution calculus for the first order linear logic. Journal of Logic, Language and Information **2** (1993) 59–83
12. Cervesato, I., Hodas, J.S., Pfenning, F.: Efficient resource management for linear logic proof search. Theoretical Computer Science **232** (2000) 133–163 Special issue on Proof Search in Type-Theoretic Languages, D. Galmiche and D. Pym, editors.
13. Harland, J., Pym, D.J.: Resource-distribution via boolean constraints. In McCune, W., ed.: Proceedings of CADE-14, Springer-Verlag LNAI 1249 (1997) 222–236
14. Girard, J.Y.: Locus solum: from the rules of logic to the logic of rules. Mathematical Structures in Computer Science **11** (2001) 301–506
15. Chaudhuri, K., Pfenning, F.: Focusing the inverse method for linear logic. Technical Report CMU-CS-05-106, Carnegie Mellon University (2005)
16. Lincoln, P., Mitchell, J.C., Scedrov, A., Shankar, N.: Decision problems for propositional linear logic. Annals of Pure and Applied Logic **56** (1992) 239–311
17. Dyckhoff, R.: Contraction-free sequent calculi for intuitionistic logic. Journal of Symbolic Logic (1992) 795–807
18. Bozzano, M., Delzanno, G., Martelli, M.: Model checking linear logic specifications. TPLP **4** (2004) 573–619