Lecture Notes on
Focusing

Oregon Summer School 2010
Proof Theory Foundations
Frank Pfenning

Lecture 4
June 17, 2010

1 Introduction

When we recast verifications as sequent proofs, we picked up a lot of redundancy in the choices we can make during proof search. Focusing is a way to eliminate much of this redundancy. On the so-called *negative* fragment of intutionistic logic, which consists of all connectives whose right rules are invertible, we can restore an isomorphism between verifications and focused proofs. If we also include positive connectives, we can do even better than verifications by further narrowing down the space of proofs with focusing.

On the negative fragment, focusing (under the name of *uniform proofs*) was described by Miller et al. [MNPS91] as a foundation for logic programming. This was greatly generalized by Andreoli [And92] in the context of proof search for classical linear logic. A recent exposition and reconstruction of focusing for a variety of logic was given by Liang and Miller [LM09] which contains further references.

Focusing also has direct applications in programming languages. For example, under the Curry-Howard isomorphism, it can be used as a foundation for functional programming with pattern matching and continuations [Zei09].

In this lecture we give a brief introduction to focusing, concentrating on the negative fragment which is of particular importance and has tractable size. At the end we generalize it to intuitionstic logic, although we do not extend the proofs of soundness and completeness, which can be found in the literature [LM09].

2 Redundancies in Proof Search

As an example, we consider the verifications of the proposition

$$(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C)).$$

In combinatory logic, this is the type of the combinator S. In the example proof, we will treat A, B, and C as atomic propositions; we have already seen that substitution in concert with local expansions allow us to obtain verifications of arbitrary instances.

The first few steps follow the general strategy of applying introductions for decomposing the goal formula.

$$\begin{array}{c} \overline{A \supset (B \supset C)\downarrow}^{\,f} \quad \overline{A \supset B\downarrow}^{\,g} \quad \overline{A\downarrow}^{\,x} \\ \vdots \\ \dfrac{\dfrac{\dfrac{C\uparrow}{A \supset C\uparrow}\supset I^x}{(A \supset B) \supset (A \supset C)\uparrow}\supset I^g}{(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))\uparrow}\supset I^f \end{array}$$

At this point we can no longer decompose the goal $C\uparrow$, so we have to work downwards from our assumptions. We start with the assumption labeled f and apply $\supset E$ twice, leaving some subgoals.

$$\begin{array}{c} \dfrac{\dfrac{\overline{A \supset (B \supset C)\downarrow}^{\,f} \quad \begin{array}{c}\vdots \\ A\uparrow\end{array}}{B \supset C\downarrow}\supset E \quad \begin{array}{c}\vdots \\ B\uparrow\end{array}}{C\downarrow}\supset E \quad \overline{A \supset B\downarrow}^{\,g} \quad \overline{A\downarrow}^{\,x} \\ \vdots \\ \dfrac{\dfrac{\dfrac{C\uparrow}{A \supset C\uparrow}\supset I^x}{(A \supset B) \supset (A \supset C)\uparrow}\supset I^g}{(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))\uparrow}\supset I^f \end{array}$$

We can verify $A\uparrow$ using $A\downarrow$, and similarly we can close the gap at the bottom by using the $\downarrow\uparrow$ conversion rule.

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{f}
\UnaryInfC{$A \supset (B \supset C)\downarrow$}
\AxiomC{}
\RightLabel{x}
\UnaryInfC{$A\downarrow$}
\RightLabel{$\downarrow\uparrow$}
\UnaryInfC{$A\uparrow$}
\RightLabel{$\supset E$}
\BinaryInfC{$B \supset C\downarrow$}
\AxiomC{}
\RightLabel{g}
\UnaryInfC{$A \supset B\downarrow$}
\AxiomC{}
\RightLabel{x}
\UnaryInfC{$A\downarrow$}
\noLine
\BinaryInfC{$\vdots$}
\noLine
\UnaryInfC{$B\uparrow$}
\RightLabel{$\supset E$}
\BinaryInfC{$C\downarrow$}
\RightLabel{$\downarrow\uparrow$}
\UnaryInfC{$C\uparrow$}
\RightLabel{$\supset I^x$}
\UnaryInfC{$A \supset C\uparrow$}
\RightLabel{$\supset I^g$}
\UnaryInfC{$(A \supset B) \supset (A \supset C)\uparrow$}
\RightLabel{$\supset I^f$}
\UnaryInfC{$(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))\uparrow$}
\end{prooftree}
$$

The final remaining gap is filled by using $\supset E$ and one conversion.

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{f}
\UnaryInfC{$A \supset (B \supset C)\downarrow$}
\AxiomC{}
\RightLabel{x}
\UnaryInfC{$A\downarrow$}
\RightLabel{$\downarrow\uparrow$}
\UnaryInfC{$A\uparrow$}
\RightLabel{$\supset E$}
\BinaryInfC{$B \supset C\downarrow$}
\AxiomC{}
\RightLabel{g}
\UnaryInfC{$A \supset B\downarrow$}
\AxiomC{}
\RightLabel{x}
\UnaryInfC{$A\downarrow$}
\RightLabel{$\downarrow\uparrow$}
\UnaryInfC{$A\uparrow$}
\RightLabel{$\supset E$}
\BinaryInfC{$B\downarrow$}
\RightLabel{$\downarrow\uparrow$}
\UnaryInfC{$B\uparrow$}
\RightLabel{$\supset E$}
\BinaryInfC{$C\downarrow$}
\RightLabel{$\downarrow\uparrow$}
\UnaryInfC{$C\uparrow$}
\RightLabel{$\supset I^x$}
\UnaryInfC{$A \supset C\uparrow$}
\RightLabel{$\supset I^g$}
\UnaryInfC{$(A \supset B) \supset (A \supset C)\uparrow$}
\RightLabel{$\supset I^f$}
\UnaryInfC{$(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))\uparrow$}
\end{prooftree}
$$

We can more compactly express this proof using the proof term notation from last lecture. We had (for $A \supset B$, $A \wedge B$ and $\top$):

$$
\begin{array}{lrcl}
\text{Normal terms} & N & ::= & \lambda x.\, N \mid \langle N_1, N_2 \rangle \mid \langle\,\rangle \mid R \\
\text{Atomic terms} & R & ::= & x \mid R\, N \mid \mathbf{fst}\, R \mid \mathbf{snd}\, R
\end{array}
$$

We can annotate the derivation we constructed above, or we can directly write out the corresponding term to obtain:

$$\lambda f.\, \lambda g.\, \lambda x.\, (f\, x)\, (g\, x) : (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))\uparrow$$

Let's try to write out the corresponding sequent proof. We must start with $\supset R$.

$$\dfrac{A \supset (B \supset C) \Longrightarrow (A \supset B) \supset (A \supset C)}{\Longrightarrow (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C)) \uparrow} \supset R$$

At this point we have a choice. We can either try to apply $\supset R$ or $\supset L$. It is a general observation that we can always eagerly apply $\supset R$ when the conclusion is an implication—we never have to consider any other rule. This is because the rule is *invertible* in the sense that if $\Gamma \Longrightarrow A \supset B$ then also $\Gamma, A \Longrightarrow B$. This invertibility follows almost immediately by weakening, identity, and cut from $\Gamma \Longrightarrow A \supset B$ and $A \supset B, A \Longrightarrow B$.

One reduction in the sequent calculus proof space then takes advantage of the *invertibility* of rules.

If we continue in this manner, after two more steps we arrive at

$$\dfrac{\dfrac{\dfrac{A \supset (B \supset C), A \supset B, A \Longrightarrow C}{A \supset (B \supset C), A \supset B \Longrightarrow A \supset C} \supset R}{A \supset (B \supset C) \Longrightarrow (A \supset B) \supset (A \supset C)} \supset R}{\Longrightarrow (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C)) \uparrow} \supset R$$

At this point we have to use an $\supset L$ rule, applied to either $A \supset (B \supset C)$ or to $A \supset B$. Both are reasonable. However, it would be nice if we could find a criterion that allows us to further reduce the choices so that only *one* of those two rules applies.

Peeking at the goal, we see that we are currently trying to prove C. The antecedent $A \supset B$ may help in that eventually, not immediately, because B is different from C. The antecedent $A \supset (B \supset C)$ mentions C in its conclusion directly, so we should focus on this assumption. If we apply the left rule there, we obtain

$$\dfrac{\dfrac{\dfrac{\dfrac{-, A \Longrightarrow A \quad -, B \supset C, A \supset B, A \Longrightarrow C}{A \supset (B \supset C), A \supset B, A \Longrightarrow C} \supset L}{A \supset (B \supset C), A \supset B \Longrightarrow A \supset C} \supset R}{A \supset (B \supset C) \Longrightarrow (A \supset B) \supset (A \supset C)} \supset R}{\Longrightarrow (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C)) \uparrow} \supset R$$

where we have elided some assumptions that are no longer relevant. At this point we have to make another choice, this time between $B \supset C$ and

$A \supset B$. But the reason we chose $A \supset (B \supset C)$ still applies, so we continue to *focus* on this assumption and obtain

$$
\dfrac{\dfrac{\dfrac{\dfrac{-, A \Longrightarrow A \qquad \dfrac{-, A \supset B, A \Longrightarrow B \quad -, C \Longrightarrow C}{-, B \supset C, A \supset B, A \Longrightarrow C}\,{\supset}L}{A \supset (B \supset C), A \supset B, A \Longrightarrow C}\,{\supset}L}{A \supset (B \supset C), A \supset B \Longrightarrow A \supset C}\,{\supset}R}{A \supset (B \supset C) \Longrightarrow (A \supset B) \supset (A \supset C)}\,{\supset}R}{\Longrightarrow (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))\uparrow}\,{\supset}R
$$

Now two subproofs can be closed off using initial sequents, and the third one after two more steps, applying the same reasoning as before (this time using antecedent $A \supset B$).

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{-, A \Longrightarrow A}\,\mathsf{init} \qquad \dfrac{\dfrac{\dfrac{}{-, A \Longrightarrow A}\,\mathsf{init} \quad \dfrac{}{-, B \Longrightarrow B}\,\mathsf{init}}{-, A \supset B, A \Longrightarrow B}\,{\supset}L \quad \dfrac{}{-, C \Longrightarrow C}\,\mathsf{init}}{-, B \supset C, A \supset B, A \Longrightarrow C}\,{\supset}L}{A \supset (B \supset C), A \supset B, A \Longrightarrow C}\,{\supset}L}{A \supset (B \supset C), A \supset B \Longrightarrow A \supset C}\,{\supset}R}{A \supset (B \supset C) \Longrightarrow (A \supset B) \supset (A \supset C)}\,{\supset}R}{\Longrightarrow (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))\uparrow}\,{\supset}R
$$

This exemplifies the second optimization we want to make in proof search. Once we have broken down the succedent into an atomic proposition, we want to restrict our choice among the antecedents to those that could directly establish the succedent. So at the first ${\supset}L$ rule we would like to rule out $A \supset B$ with a minimum of effort. This optimization is called *focusing*, which is also the term for the overall strategy and therefore slightly ambiguous, but we will use it nevertheless.

3 Focusing in the Negative Fragment

The fragment with implication, conjunction, and truth is called the *negative fragment*. It has the property that all of its right rules are invertible, while the positive connectives (disjunction and falsehood) discussed in Section 7 have invertible left rules.

In proof theory we try to capture the essence of strategies with inference systems as much as possible. We have already seen this in the definition of truth, verifications, and the sequent calculus. So here we have a new judgment $\Gamma \longrightarrow A$ expressing that A has a focused proof. To capture inversion, there is exactly one rule for each possible succedent.

$$\frac{\Gamma, A \longrightarrow B}{\Gamma \longrightarrow A \supset B} \supset R \qquad \frac{\Gamma \longrightarrow A \quad \Gamma \longrightarrow B}{\Gamma \longrightarrow A \wedge B} \wedge R \qquad \frac{}{\Gamma \longrightarrow \top} \top R$$

When the succedent is atomic, we *focus* on one of the antecedents, actually copying from the ambient context to a special position, denoted by $[A]$. This form of sequent is written as $\Gamma; [A] \longrightarrow P$ and means that P has a proof, focusing on A. Focus is inherited by the premises so that we are forced to continue to use the subformulas of the focus formula until we have reduced it to the atomic case.

$$\frac{\Gamma; [A] \longrightarrow P \quad A \in \Gamma}{\Gamma \longrightarrow P} \,\mathsf{focus}$$

$$\frac{\Gamma \longrightarrow A \quad \Gamma; [B] \longrightarrow P}{\Gamma; [A \supset B] \longrightarrow P} \supset L \qquad \frac{\Gamma; [A] \longrightarrow P}{\Gamma; [A \wedge B] \longrightarrow P} \wedge L_1 \qquad \frac{\Gamma; [B] \longrightarrow P}{\Gamma; [A \wedge B] \longrightarrow P} \wedge L_2$$

There is no $\top L$ rule, just as there is none in the sequent calculus. Note that in these rules we do not copy the focus formula to the premises. We can afford to do that because we copy the original formula from Γ in the rule focus, so a copy of A is retained. In fact, the judgment $\Gamma; [A] \longrightarrow P$ is hypothetical in Γ (which is persistent: any antecedent will remain available in the remainder of the proof), but not hypothetical in $[A]$. Indeed, for those who know about linear logic, it is *linear* in $[A]$: any use of $[A]$ will consume it in the proof, and it must be consumed.

Once we reduce the focus formula to be atomic, we either succeed or fail, depending on the succedent.

$$\frac{}{\Gamma; [P] \longrightarrow P} \,\mathsf{init} \qquad \begin{array}{c}\text{no rule for } Q \neq P \\ \Gamma; [Q] \longrightarrow P\end{array}$$

In order to limit the nondeterminism, there is a small amount of extra control knowledge we apply. When using the focused $\supset L$ rule we continue to use the *right* premise first, continuing our focus in $[B]$ which inherits its focus from $[A \supset B]$. This means that focusing fails in only a few steps, without any real search, when we focus on an assumption that does not help us to establish P right away. It is this observation from which focusing draws a big part of its utility.

We now revisit the earlier example in focused form. The initial sequence of inversion steps is entirely forced.

$$\dfrac{\dfrac{\dfrac{A \supset (B \supset C), A \supset B, A \longrightarrow C}{A \supset (B \supset C), A \supset B \longrightarrow A \supset C} \supset R}{A \supset (B \supset C) \longrightarrow (A \supset B) \supset (A \supset C)} \supset R}{\longrightarrow (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))} \supset R$$

Now we can focus on any of the three antecedents. Focusing on A or $A \supset B$ will fail quickly, after one or two steps, so we focus on $A \supset (B \supset C)$. We pursue the forced branch of the proof, with continued focus, leaving open subgoals on the side. We abbreviate $(A \supset (B \supset C), A \supset B, A) = \Gamma_0$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Gamma_0 \longrightarrow A \qquad \dfrac{\Gamma_0 \longrightarrow B \qquad \dfrac{}{\Gamma_0; [C] \longrightarrow C} \mathsf{init}}{\Gamma_0; [B \supset C] \longrightarrow C} \supset L}{\Gamma_0; [A \supset (B \supset C)] \longrightarrow C} \supset L}{A \supset (B \supset C), A \supset B, A \longrightarrow C} \mathsf{focus}}{A \supset (B \supset C), A \supset B \longrightarrow A \supset C} \supset R}{A \supset (B \supset C) \longrightarrow (A \supset B) \supset (A \supset C)} \supset R}{\longrightarrow (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))} \supset R$$

By focusing on $A \supset (B \supset C)$ we have replaced the goal of proving C by the subgoals of proving A and B.

For the first subgoal, we have to focus on A, since focus in $A \supset (B \supset C)$ or $A \supset B$ will fail quickly. For the second we focus on $A \supset B$ for similar reasons.

\begin{prooftree}
\AxiomC{}
\RightLabel{init}
\UnaryInfC{$\Gamma_0; [A] \longrightarrow A$}
\RightLabel{focus}
\UnaryInfC{$\Gamma_0 \longrightarrow A$}
\AxiomC{$\Gamma_0 \longrightarrow A$}
\AxiomC{}
\RightLabel{init}
\UnaryInfC{$\Gamma_0; [B] \longrightarrow B$}
\RightLabel{$\supset L$}
\BinaryInfC{$\Gamma_0; [A \supset B] \longrightarrow B$}
\RightLabel{focus}
\UnaryInfC{$\Gamma_0 \longrightarrow B$}
\AxiomC{}
\RightLabel{init}
\UnaryInfC{$\Gamma_0; [C] \longrightarrow C$}
\RightLabel{$\supset L$}
\BinaryInfC{$\Gamma_0; [B \supset C] \longrightarrow C$}
\RightLabel{$\supset L$}
\BinaryInfC{$\Gamma_0; [A \supset (B \supset C)] \longrightarrow C$}
\RightLabel{focus}
\UnaryInfC{$A \supset (B \supset C), A \supset B, A \longrightarrow C$}
\RightLabel{$\supset R$}
\UnaryInfC{$A \supset (B \supset C), A \supset B \longrightarrow A \supset C$}
\RightLabel{$\supset R$}
\UnaryInfC{$A \supset (B \supset C) \longrightarrow (A \supset B) \supset (A \supset C)$}
\RightLabel{$\supset R$}
\UnaryInfC{$\longrightarrow (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$}
\end{prooftree}

The last open subgoal can again be proven by focusing on A as before.

\begin{prooftree}
\AxiomC{}
\RightLabel{init}
\UnaryInfC{$\Gamma_0; [A] \longrightarrow A$}
\RightLabel{focus}
\UnaryInfC{$\Gamma_0 \longrightarrow A$}
\AxiomC{}
\RightLabel{init}
\UnaryInfC{$\Gamma_0; [A] \longrightarrow A$}
\RightLabel{focus}
\UnaryInfC{$\Gamma_0 \longrightarrow A$}
\AxiomC{}
\RightLabel{init}
\UnaryInfC{$\Gamma_0; [B] \longrightarrow B$}
\RightLabel{$\supset L$}
\BinaryInfC{$\Gamma_0; [A \supset B] \longrightarrow B$}
\RightLabel{focus}
\UnaryInfC{$\Gamma_0 \longrightarrow B$}
\AxiomC{}
\RightLabel{init}
\UnaryInfC{$\Gamma_0; [C] \longrightarrow C$}
\RightLabel{$\supset L$}
\BinaryInfC{$\Gamma_0; [B \supset C] \longrightarrow C$}
\RightLabel{$\supset L$}
\BinaryInfC{$\Gamma_0; [A \supset (B \supset C)] \longrightarrow C$}
\RightLabel{focus}
\UnaryInfC{$A \supset (B \supset C), A \supset B, A \longrightarrow C$}
\RightLabel{$\supset R$}
\UnaryInfC{$A \supset (B \supset C), A \supset B \longrightarrow A \supset C$}
\RightLabel{$\supset R$}
\UnaryInfC{$A \supset (B \supset C) \longrightarrow (A \supset B) \supset (A \supset C)$}
\RightLabel{$\supset R$}
\UnaryInfC{$\longrightarrow (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$}
\end{prooftree}

Of course, it is incumbent upon us to show that this system is strong enough to show that it can find a focused proof for every proposition that has a verification. We will do this in the section after the next one.

4 Big-Step Derived Rules of Inference

The[1] focused rules eliminate a lot of nondeterminism, but this may difficult to see since there is a lot of bureaucracy in the rules. Since our hand is completely forced until we have an atomic conclusion, we can also derive rules of inference which capture what would happen if we did focus on an assumption. The complete collection of these derived rules is complete for verifications.

In the particular running example of this lecture

$$(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$$

we have the following formulas that can appear as a hypothesis in a proof and could therefore be focused on: $A \supset (B \supset C)$, $A \supset B$, and A. Let's set up a situation with an unknown context Γ and goal P and trace through a phase of focusing. First, $A \supset (B \supset C)$.

$$\dfrac{\dfrac{\Gamma \longrightarrow A \qquad \dfrac{\Gamma \longrightarrow B \qquad \dfrac{C = P}{\Gamma; [C] \longrightarrow P}\,\mathsf{init}}{\Gamma; [B \supset C] \longrightarrow P}\,{\supset}L}{\Gamma; [A \supset (B \supset C)] \longrightarrow P}\,{\supset}L}{\Gamma \longrightarrow P}\,\mathsf{focus}$$

Note that all inference rules are forced by the focusing discipline, as is the condition that $C = P$. Substituting C for P we see that this hypothetical derivation establishes the derived rule

$$\dfrac{\Gamma \longrightarrow A \quad \Gamma \longrightarrow B}{\Gamma \longrightarrow C}\,[A \supset (B \supset C)]$$

and that all other attempts at focusing on $A \supset (B \supset C)$ will fail.

Focusing on $A \supset B$ and A, respectively, yields the following two additional rules.

$$\dfrac{\Gamma \longrightarrow A}{\Gamma \longrightarrow B}\,[A \supset B] \qquad\qquad \dfrac{}{\Gamma \longrightarrow A}\,[A]$$

After applying our initial inversion steps we arrive at $\Gamma_0 \longrightarrow C$ where $\Gamma_0 = (A \supset (B \supset C), A \supset B, A)$ as before. Now we can complete the proof

[1] This bonus section did not fit into the allotted lecture time.

using only the derived rules.

$$\dfrac{\dfrac{}{\Gamma_0 \longrightarrow A}[A] \quad \dfrac{\dfrac{}{\Gamma_0 \longrightarrow A}[A]}{\Gamma_0 \longrightarrow B}[A \supset B]}{\Gamma_0 \longrightarrow C}[A \supset (B \supset C)]$$

In this application of focusing we just use one form of sequent, $\Gamma \longrightarrow P$, since the focusing sequent $\Gamma; [A] \longrightarrow P$ is completely folded into the derived rules, as are the inversion steps to be applied when the succedent is not atomic.

We will not formally describe the process of converting formulas to rules, but it is a very general and powerful tools in building theorem provers for non-classical logics (see, for example, [MP09]).

5 Completeness of Focusing

It is very easy to see that focusing is *sound* with respect to the sequent calculus. We merely map $\Gamma \longrightarrow A$ to $\Gamma \Longrightarrow A$ and $\Gamma; [A] \longrightarrow C$ to $\Gamma, A \Longrightarrow C$. Then all the inference rules for the focusing system turn into regular sequent calculus rules, where we have to weaken in the left rule to match the premises exactly. For the focus rule we note that contraction is admissible in the sequent calculus (see Lecture 3). From there, by a prior theorem, we can go to verifications.

The more difficult property is completeness with respect to the sequent calculus. Since our ultimate goal is the relation to verification, we go directly from verifications to focused proofs. We use exactly the same structure as in the proof of the completeness of the sequent calculus, but force the left rules in focusing.

Theorem 1 (From Verifications to Focused Proofs)

(i) If $\Gamma^{\downarrow} \vdash A\uparrow$ *then* $\Gamma \longrightarrow A$

(ii) If $\Gamma^{\downarrow} \vdash A\downarrow$ *and* $\Gamma; [A] \longrightarrow P$ *then* $\Gamma \longrightarrow P$ *for any* P.

Proof: By induction on the structure of the given verification. We only show a few cases; the others are similar, or follow the one for the sequent calculus.

Case:

$$\frac{\Gamma^{\downarrow} \vdash P\!\downarrow}{\Gamma^{\downarrow} \vdash P\!\uparrow}\ {\downarrow\uparrow}$$

$\Gamma; [P] \longrightarrow P$	By rule init
$\Gamma \longrightarrow P$	By i.h.(ii)

Case:

$$\frac{A\!\downarrow \in \Gamma^{\downarrow}}{\Gamma^{\downarrow} \vdash A\!\downarrow}\ \mathsf{hyp}$$

$\Gamma; [A] \longrightarrow P$	Assumption
$\Gamma \longrightarrow P$	By rule focus (since $A \in \Gamma$)

Case:

$$\frac{\Gamma^{\downarrow} \vdash B \supset A\!\downarrow \quad \Gamma^{\downarrow} \vdash B\!\uparrow}{\Gamma^{\downarrow} \vdash A\!\downarrow}\ {\supset}E$$

$\Gamma; [A] \longrightarrow P$	Assumption
$\Gamma \longrightarrow B$	By i.h.(i)
$\Gamma; [B \supset A] \longrightarrow P$	By rule $\supset L$
$\Gamma \longrightarrow P$	By i.h.(ii)

□

6 Proof Terms

The proof terms we assign to focused proofs come from the so-called *spine calculus* [CP03], an extension of an earlier proof term assignment by Herbelin [Her95]. For the negative fragment, this allows us to establish a *bijection* between verifications and focusing proofs. The advantage of focusing proofs is that all proofs are constructed bottom-up, while verifications go in two directions, which simplifies the construction of a proof search procedure and proofs of properties of the calculus.

We create a new class of terms U and spines S. We use the following syntax:

$$\begin{array}{lcl} \text{Terms} & U & ::= & \lambda x.\, U \mid \langle U_1, U_2 \rangle \mid \langle\,\rangle \mid x \cdot S \\ \text{Spines} & S & ::= & (U \,;\, S) \mid \pi_1\, S \mid \pi_2\, S \mid (\,) \end{array}$$

The meaning of the constructs is probably best understood by the annotation of the proof rules, which double as typing rules under the usual Curry-Howard correspondence. The two annotated judgments are

$\Gamma \longrightarrow U : A$	U is a focused term of type A
$\Gamma; [A] \longrightarrow S : P$	S is a spine mapping a head of type A to a term of type P

$$\frac{\Gamma, x{:}A \longrightarrow U : B}{\Gamma \longrightarrow \lambda x.\, U : A \supset B}\ \supset R \qquad \frac{\Gamma \longrightarrow U : A \quad \Gamma \longrightarrow V : B}{\Gamma \longrightarrow \langle U, V \rangle : A \wedge B}\ \wedge R \qquad \frac{}{\Gamma \longrightarrow \langle\,\rangle : \top}\ \top R$$

$$\frac{\Gamma; [A] \longrightarrow S : P \quad x{:}A \in \Gamma}{\Gamma \longrightarrow x \cdot S : P}\ \mathsf{focus}$$

$$\frac{\Gamma \longrightarrow U : A \quad \Gamma; [B] \longrightarrow S : P}{\Gamma; [A \supset B] \longrightarrow (U\ ;\ S) : P}\ \supset L$$

$$\frac{\Gamma; [A] \longrightarrow S : P}{\Gamma; [A \wedge B] \longrightarrow \pi_1\, S : P}\ \wedge L_1 \qquad \frac{\Gamma; [B] \longrightarrow S : P}{\Gamma; [A \wedge B] \longrightarrow \pi_2\, S : P}\ \wedge L_2$$

$$\frac{}{\Gamma; [P] \longrightarrow (\,) : P}\ \mathsf{init} \qquad \begin{array}{c}\text{no rule for } Q \neq P \\ \Gamma; [Q] \longrightarrow - : P\end{array}$$

On an applicative term, the translation from verifications to focused terms is a re-association: the term $((x\, N_1)\, N_2)$ becomes $x \cdot (N_1\ ;\ (N_2\ ;\ (\,)))$. Projections are mapped to spine operators and listed inside-out, so that **fst** (**snd** x) becomes $x \cdot (\pi_2\,(\pi_1\,(\,)))$.

We can now exhibit the bijection between (well-typed) terms in the two forms. The fact that we have to re-associate the arguments suggests using an accumulator argument in the translation of atomic terms, while working compositionally with normal terms. We have $(N)^* = U$ and $R \mathbin{@} S = U$ where S accumulates a spine from back to front.

$$\begin{array}{lcl}
(\lambda x.\, N)^* & = & \lambda x.\, N^* \\
\langle N_1, N_2 \rangle^* & = & \langle N_1^*, N_2^* \rangle \\
\langle\,\rangle^* & = & \langle\,\rangle \\
R^* & = & R \mathbin{@} (\,) \\[1ex]
(R\, N) \mathbin{@} S & = & R \mathbin{@} (N^*\ ;\ S) \\
(\mathbf{fst}\, R) \mathbin{@} S & = & R \mathbin{@} (\pi_1\, S) \\
(\mathbf{snd}\, R) \mathbin{@} S & = & R \mathbin{@} (\pi_2\, S) \\
x \mathbin{@} S & = & x \cdot S
\end{array}$$

It is now easy to see how to define the reverse translation: if the focused terms has the form $x \cdot S$ we translate it to $x \mathbin{@} S$ where now x initializes an accumulator R. When S has been reduced to $(\,)$ we return R.

7 Polarized Focusing

When we try to generalize to arbitrary propositions, we see that disjunction and falsehood are essentially different from implication, conjunction, and truth. This is because their left rules in the sequent calculus are invertible, rather than their right rules. A related symptom is that their elimination rules in natural deduction have to use an auxiliary formula C.

Both inversion and focusing phases of proof search chain together similar inferences (invertible or focused), no matter what the connective. It is therefore convenient to *polarize* the formula by combining runs of connectives that are positive or negative, with explicit coercions between the runs. These coercions, written $\uparrow$ and $\downarrow$ are called *shift operators*. Applying this idea we obtain:

$$\begin{array}{lll} \text{Negative propositions} & A^- & ::= \;\; P^- \mid A^+ \supset A^- \mid A^- \wedge A^- \mid \top \mid {\uparrow}A^+ \\ \text{Positive propositions} & A^+ & ::= \;\; A^+ \vee A^+ \mid \bot \mid {\downarrow}A^- \end{array}$$

Note that the left-hand side of an implication is positive because it will end up on the other side of the sequent from the implication itself.

We can extend this further by noticing that the following new left rules for conjunction and truth are in fact invertible.

$$\frac{\Gamma, A, B \Longrightarrow C}{\Gamma, A \wedge B \Longrightarrow C}\ \wedge L \qquad \frac{\Gamma \Longrightarrow C}{\Gamma, \top \Longrightarrow C}\ \top L$$

We have written P^- for a propositional variable P stands for a negative proposition, which suggests we should also have propositional variables P^+ that stand for positive propositions. Taking both of these observations into account we obtain:

$$\begin{array}{lll} \text{Negative propositions} & A^- & ::= \;\; P^- \mid A^+ \supset A^- \mid A^- \wedge A^- \mid \top \mid {\uparrow}A^+ \\ \text{Positive propositions} & A^+ & ::= \;\; P^+ \mid A^+ \vee A^+ \mid \bot \mid A^+ \wedge A^+ \mid \top \mid {\downarrow}A^- \end{array}$$

A *polarized sequent* has the form $\Gamma^+ \longrightarrow A^-$, that is, all hypotheses are positive and the conclusion is always negative. The invertible rules proceed directly on such a polarized sequent.

$$\frac{\Gamma^+, A^+ \longrightarrow B^-}{\Gamma^+ \longrightarrow A^+ \supset B^-} \supset R \quad \frac{\Gamma^+ \longrightarrow A^- \quad \Gamma^+ \longrightarrow A^-}{\Gamma^+ \longrightarrow A^- \wedge B^-} \wedge R \quad \frac{}{\Gamma^+ \longrightarrow \top} \top R$$

$$\frac{\Gamma^+, A^+ \longrightarrow C^- \quad \Gamma^+, B^+ \longrightarrow C^-}{\Gamma^+, A^+ \vee B^+ \longrightarrow C^-} \vee L \quad \frac{}{\Gamma^+, \bot \longrightarrow C^-} \bot L$$

$$\frac{\Gamma^+, A^+, B^+ \longrightarrow C^-}{\Gamma^+, A^+ \wedge B^+ \longrightarrow C^-} \wedge L \quad \frac{\Gamma^+ \longrightarrow C^-}{\Gamma^+, \top \longrightarrow C^-} \top L$$

The intent is to apply all the invertible rules on polarized sequents in a don't-care nondeterministic manner. Because all these invertible rules decrease the complexity of the sequent by eliminating connectives this phase will have to come to completion. At this point, the succedent will be either P^- or $\uparrow C^+$ while each antecedent will be either P^+ or $\downarrow C^-$. We can now focus either on an antecedent or on the succedent of the right form.

$$\frac{\Gamma^+; [A^-] \longrightarrow C^- \quad \text{for } \downarrow A^- \in \Gamma^+}{\Gamma^+ \longrightarrow C^-} \downarrow L \qquad \frac{\Gamma^+ \longrightarrow [C^+]}{\Gamma^+ \longrightarrow \uparrow C^+} \uparrow R$$

As written, this rule enforces only *weak focusing*, where inversion is not forced but can be applied in a discretionary way. If we enforce the side condition mentioned above, we have *full focusing*.

We can see that these rules require two different focusing judgments, one on the left, where the single focus formula must be negative, or on the right, where the focus formula must be positive. They are defined by the following rules:

$$\frac{\Gamma^+ \longrightarrow [A^+] \quad \Gamma^+, [B^-] \longrightarrow C^-}{\Gamma^+; [A^+ \supset B^-] \longrightarrow C^-} \supset L$$

$$\frac{\Gamma^+; [A^-] \longrightarrow C^-}{\Gamma^+; [A^- \wedge B^-] \longrightarrow C^-} \wedge L_1 \quad \frac{\Gamma^+; [B^-] \longrightarrow C^-}{\Gamma^+; [A^- \wedge B^-] \longrightarrow C^-} \wedge L_2 \quad \text{no } \top L \text{ in focus}$$

$$\frac{\Gamma^+ \longrightarrow [A^+]}{\Gamma^+ \longrightarrow [A^+ \vee B^+]} \vee R_1 \quad \frac{\Gamma^+ \longrightarrow [B^+]}{\Gamma^+ \longrightarrow [A^+ \vee B^+]} \vee R_2 \quad \text{no } \bot R$$

$$\frac{\Gamma^+ \longrightarrow [A^+] \quad \Gamma^+ \longrightarrow [B^+]}{\Gamma^+ \longrightarrow [A^+ \wedge B^+]} \wedge R \quad \frac{}{\Gamma^+ \longrightarrow [\top]} \top R$$

$$\frac{}{\Gamma^+; [P^-] \longrightarrow P^-} \mathsf{init} \quad \begin{array}{c}\text{no rule for } Q^- \neq P^- \\ \Gamma^+; [Q^-] \longrightarrow P^-\end{array}$$

$$\frac{P^+ \in \Gamma^+}{\Gamma^+ \longrightarrow [P^+]} \mathsf{init} \quad \begin{array}{c}\text{no rule for } P^+ \notin \Gamma^+ \\ \Gamma^+ \longrightarrow [P^+]\end{array}$$

Once the focusing phase is complete, we must encounter shift operators, at which point we return to a regular focused sequent.

$$\frac{\Gamma^+, A^+ \longrightarrow C^-}{\Gamma^+; [\uparrow A^+] \longrightarrow C^-} \uparrow L \quad \frac{\Gamma^+ \longrightarrow A^-}{\Gamma^+ \longrightarrow [\downarrow A^-]} \downarrow R$$

This inference system does not tell the full story of proof search. One important point is that the invertible rules are applied in an arbitrary fashion, but one does not backtrack over the choices. This can be enforced by creating an *ordered context* (see, for example, Andreoli [And92] or Liang and Miller [LM09]) but we will not do so here.

If we start with an ordinary, unpolarized proposition there are a number of ways to insert the shift operators $\uparrow$ and $\downarrow$ so that we arrive at a polarized one. The search space will be different, taking bigger or smaller steps, but each annotation is sound and complete as long as erasure of shifts leads back to the original (unpolarized) formula. On one extreme, we can recover the usual single-step sequent calculus by shifting for each subformula so that each phase consists of breaking down just one connective. On the other extreme is a which inserts the minimal number of shifts in the proof of a proposition A (which therefore must be negative at the top level). Before we apply the translation, we must assign polarities to all the atoms, so that their target will be consistently polarized. This apparently innocent choice on the polarity of atoms can have drastic consequences in the search space. For example, on the Horn fragment we can recover either bottom-up logic programming ot top-down logic programming purely by chosing an appropriate polarity for atoms [CPP08].

References

[And92] Jean-Marc Andreoli. Logic programming with focusing proofs in linear logic. *Journal of Logic and Computation*, 2(3):197–347, 1992.

[CP03] Iliano Cervesato and Frank Pfenning. A linear spine calculus. *Journal of Logic and Computation*, 13(5):639–688, 2003.

[CPP08] Kaustuv Chaudhuri, Frank Pfenning, and Greg Price. A logical characterization of forward and backward chaining in the inverse method. *Journal of Automated Reasoning*, 40(2–3):133–177, 2008. Special issue with selected papers from IJCAR 2006.

[Her95] Hugo Herbelin. *Séquents qu'on calcule*. PhD thesis, University Paris 7, January 1995.

[LM09] Chuck Liang and Dale Miller. Focusing and polarization in linear, intuitionistic, and classical logics. *Theoretical Computer Science*, 410(46):4747–4768, November 2009.

[MNPS91] Dale Miller, Gopalan Nadathur, Frank Pfenning, and Andre Scedrov. Uniform proofs as a foundation for logic programming. *Annals of Pure and Applied Logic*, 51:125–157, 1991.

[MP09] Sean McLaughlin and Frank Pfenning. Efficient intuitionistic theorem proving with the polarized inverse method. In R.A.Schmidt, editor, *Proceedings of the 22nd International Conference on Automated Deduction (CADE-22)*, pages 230–244, Montreal, Canada, August 2009. Springer LNCS 5663.

[Zei09] Noam Zeilberger. *The Logical Basis of Evaluation Order and Pattern-Matching*. PhD thesis, Department of Computer Science, Carnegie Mellon University, May 2009. Available as Technical Report CMU-CS-09-122.