Lecture Notes on Goal Stacks

15-317: Constructive Logic
Frank Pfenning

Lecture 20
Tuesday, April 4, 2023

1 Introduction

We have emphasized the central role of high level rule descriptions: on one hand they are abstract enough to allow formal proof of correctness, and on the other hand they are concrete enough to support an implementation (be it directly in logic programming or indirectly in functional programming). Today, we'll encounter another example of this and also practice rule induction.

When we looked at proof reduction as a basic mechanism for computation we carefully developed a notion of small step reduction from the local reduction delivered to us from proof theory such that it satisfied preservation, progress, and determinism. The metacircular interpreter from Lecture 19 does not quite satisfy that because the details of proof construction in the object logic still depends on details of proof construction in the metalogic. And both of them are Horn clauses! The choices left are:

1. Subgoal ordering: in which order do we solve $G_1 \wedge G_2$? The informal strategy is to derive G_1 before G_2.
2. Backtracking: in which order do we try the program clauses? The informal strategy is to try them in the order the are presented.
3. Unification: how do we avoid guessing terms when quantifiers are instantiated? The informal strategy is to find a most general solution to the equations imposed by matching atomic propositions.

In this lecture we tackle subgoal ordering; in the next lecture we take a look at unification.

2 Continuations as Stacks of Goals

Let's recall the rules for backward chaining from the last lecture in Figure 1. Only one of these rules, namely $\wedge R$ has more than premise. We have added $\top$ as another possible goal because that turns out to be convenient for today's lecture. The way we have been writing

$$\begin{array}{lll}
\text{Horn clauses} & D & ::= \ \forall x.\, D \mid G \supset P \mid P \\
\text{Goals} & G & ::= \ P \mid G_1 \wedge G_2 \mid \top \mid \exists x.\, G(x)
\end{array}$$

Choice $\Gamma \xrightarrow{\mathsf{C}} P$

$$\frac{D \in \Gamma \quad \Gamma, [D] \xrightarrow{\mathsf{FL}} P}{\Gamma \xrightarrow{\mathsf{C}} P}\ \mathsf{FLC}$$

Left Focus $\Gamma, [D] \xrightarrow{\mathsf{FL}} P$

$$\frac{}{\Gamma, [P] \xrightarrow{\mathsf{FL}} P}\ \mathsf{id} \qquad \frac{\Gamma \xrightarrow{\mathsf{FR}} [G]}{\Gamma, [G \supset P] \xrightarrow{\mathsf{FL}} P}\ {\supset}L \qquad \frac{\Gamma, [D(t)] \xrightarrow{\mathsf{FL}} P}{\Gamma, [\forall x.\, D(x)] \xrightarrow{\mathsf{FL}} P}\ \forall L$$

Right Focus $\Gamma \xrightarrow{\mathsf{FR}} [G]$**.**

$$\frac{\Gamma \xrightarrow{\mathsf{FR}} [G_1] \quad \Gamma \xrightarrow{\mathsf{FR}} [G_2]}{\Gamma \xrightarrow{\mathsf{FR}} [G_1 \wedge G_2]}\ \wedge R \qquad \frac{}{\Gamma \xrightarrow{\mathsf{FR}} [\top]}\ \top R \qquad \frac{\Gamma \xrightarrow{\mathsf{FR}} [G(t)]}{\Gamma \xrightarrow{\mathsf{FR}} [\exists x.\, G(x)]}\ \exists R \qquad \frac{\Gamma \xrightarrow{\mathsf{C}} P}{\Gamma \xrightarrow{\mathsf{FR}} [P]}\ \mathsf{CFR}$$

Figure 1: Backward chaining for Horn clauses

and interpreting these rules we have a strong intuition that the first premise should be derived first and then the second, but the rule itself does not express that.

So we would like to transform these rules into an equivalent set so that each rule has at most one premise. That would then definitively resolve the question. The main idea is to revise $\wedge R$ to have one premise in which we derive G_1. But that means we have to "remember" that G_2 still has to be solved after G_1 is done. We accomplish this by pushing G_2 only a stack of subgoals. We'll reuse the an ordered collection of propositions Ω for this purpose since it makes the ordered nature as explicit as possible.

$$\text{Goal stacks} \quad \Omega \quad ::= \quad G \cdot \Omega \mid \epsilon$$

The new judgment for focus on the right is $\Gamma \xrightarrow{\mathsf{FR}} [G] \cdot \Omega$.

$$\frac{\Gamma \xrightarrow{\mathsf{FR}} [G_1] \cdot (G_2 \cdot \Omega)}{\Gamma \xrightarrow{\mathsf{FR}} [G_1 \wedge G_2] \cdot \Omega} \wedge R$$

This rule looks almost like a structural rule, but it does in fact eliminate the conjunction so it is a new version of $\wedge R$. When we succeed (let's consider the $\top$ case), we still have to go back and focus on the first thing in Ω. Let's call this *choice right* (written $\Gamma \xrightarrow{\mathsf{CR}} \Omega$), even though it is really a transition judgment.

$$\frac{\Gamma \xrightarrow{\mathsf{CR}} \Omega}{\Gamma \xrightarrow{\mathsf{FR}} [\top] \cdot \Omega} \top R$$

$$\frac{\Gamma \xrightarrow{\mathsf{FR}} [G] \cdot \Omega}{\Gamma \xrightarrow{\mathsf{CR}} G \cdot \Omega} \mathsf{FRCR} \qquad \frac{}{\Gamma \xrightarrow{\mathsf{CR}} \epsilon} \mathsf{CR}$$

When the goal stack is empty, we succeed. Otherwise, we focus on the goal G on top of the stack. Interesting here is that $\top R$, usually the end of the derivation, has a premise in which the goal stack is consulted to see if there are any remaining unsolved subgoals.

Now we have to propagate the goal stack through the remaining rules and examine the consequences. Sometimes this process of rule engineering succeeds straightforwardly (like here), and sometimes it points to problems and judgments may have to revised further to accommodate the interaction of multiple features.

Finishing right focus:

$$\frac{\Gamma \xrightarrow{\mathsf{FR}} [G(t)] \cdot \Omega}{\Gamma \xrightarrow{\mathsf{FR}} [\exists x.\, G(x)] \cdot \Omega} \exists R \quad \frac{\Gamma \xrightarrow{\mathsf{CL}} P \cdot \Omega}{\Gamma \xrightarrow{\mathsf{FR}} [P] \cdot \Omega} \mathsf{CFR}$$

We see that the choice judgment (which we have renamed to *choice left* $\Gamma \xrightarrow{\mathsf{CL}} P \cdot \Omega$) just carries an atom on the right, followed by the remaining goal stack Ω. However, it doesn't change in any significant way, threading through the goal stack.

$$\frac{D \in \Gamma \quad \Gamma, [D] \xrightarrow{\mathsf{FL}} P \cdot \Omega}{\Gamma \xrightarrow{\mathsf{CL}} P \cdot \Omega} \mathsf{FLCL}$$

Left focus is interesting: the identity rule now has a premise because we still have to solve the goals remaining in the goal stack! In contrast, the $\supset L$ rule continues focus on G as before, carrying Ω along. That's because we want to solve the immediate subgoal G before any other potential goals in the goal stack.

$$\frac{\Gamma \xrightarrow{\mathsf{CR}} \Omega}{\Gamma, [P] \xrightarrow{\mathsf{FL}} P \cdot \Omega}\ \mathsf{id} \qquad \frac{\Gamma \xrightarrow{\mathsf{FR}} [G] \cdot \Omega}{\Gamma, [G \supset P] \xrightarrow{\mathsf{FL}} P \cdot \Omega}\ {\supset}L \qquad \frac{\Gamma, [D(t)] \xrightarrow{\mathsf{FL}} P \cdot \Omega}{\Gamma, [\forall x.\, D(x)] \xrightarrow{\mathsf{FL}} P \cdot \Omega}\ \forall L$$

Scanning the rules on this and the previous page, we see have achieved our goal: all rules have at most one premise. Subgoal ordering is now explicit. In fact, all the rules have *exactly* one premise, except for CR which is the only one that can complete a derivation.

We view the rule FLCL which picks among the clauses in Γ as having a single premise, although there are many choices on how to apply this rules (one possibility for each $D \in \Gamma$).

After completing this rule design we can see that the goal stack Ω acts as a *continuation*, that is, a data structure that prescribes what to do after the current goal has been successfully solved. Because of that, we call it a *success continuation*.

3 Soundness of Goal Stacks

As just noted, the rules achieve our goal to make subgoal ordering explicit. But are they "correct"? For that we need to relate them back to our original backchaining rules. If you haven't done this kind of proof before (or even if you have), this is surprisingly tricky to get just right. Usually, the more straightforward direction is from the new (generally more complex) version to the earlier (generally simpler) version of the rules. Often this is called *soundness*, because the new rules are sound with respect to the earlier ones.

In this case, the key is that if we have $\Gamma \xrightarrow{\mathsf{CR}} \Omega$ then we can prove all the goals in Ω. For this we have to map the stack back to a proposition.

$$\begin{array}{lcl} \bigwedge(\epsilon) & = & \top \\ \bigwedge(G \cdot \Omega) & = & G \wedge \bigwedge \Omega \end{array}$$

We also notice that we have four judgments, so our theorem has four parts. We encourage you to try to formulate your own theorem and induction before consulting our proof.

Theorem 1 (Soundness of Goal Stacks)

(1) If $\Gamma \xrightarrow{\mathsf{CR}} \Omega$ *then* $\Gamma \xrightarrow{\mathsf{FR}} [\bigwedge \Omega]$

(2) If $\Gamma \xrightarrow{\mathsf{CL}} P \cdot \Omega$ *then* $\Gamma \xrightarrow{\mathsf{C}} P$ *and* $\Gamma \xrightarrow{\mathsf{FR}} [\bigwedge \Omega]$

(3) If $\Gamma \xrightarrow{\mathsf{FR}} [G] \cdot \Omega$ *then* $\Gamma \xrightarrow{\mathsf{FR}} [G]$ *and* $\Gamma \xrightarrow{\mathsf{FR}} [\bigwedge \Omega]$

(4) If $\Gamma, [D] \xrightarrow{\mathsf{FL}} P \cdot \Omega$ *then* $\Gamma, [D] \xrightarrow{\mathsf{FL}} P$ *and* $\Gamma \xrightarrow{\mathsf{FR}} [\bigwedge \Omega]$

Proof: By simultaneous rule induction the given derivation $\mathcal{D}$ in parts (1)–(4). We assume we are given a derivation $\mathcal{D}$ and have to construction two derivations in each part. We show a few cases; the others are similar.

Case (1):

$$\mathcal{D} \quad = \dfrac{\begin{array}{c}\mathcal{D}' \\ \Gamma \xrightarrow{\mathsf{FR}} [G]\cdot\Omega\end{array}}{\Gamma \xrightarrow{\mathsf{CR}} G\cdot\Omega}\ \mathsf{FRCR}$$

Then we construct

$$\dfrac{\begin{array}{c}\mathsf{IH}_3(\mathcal{D}') \\ \Gamma \xrightarrow{\mathsf{FR}} [G]\end{array} \quad \begin{array}{c}\mathsf{IH}_3(\mathcal{D}') \\ \Gamma \xrightarrow{\mathsf{FR}} [\bigwedge\Omega]\end{array}}{\Gamma \xrightarrow{\mathsf{FR}} [G \land \bigwedge\Omega]}\ {\land}R$$

Case (3):

$$\mathcal{D} = \quad \dfrac{\begin{array}{c}\mathcal{D}_1 \\ \Gamma \xrightarrow{\mathsf{FR}} [G_1]\cdot(G_2\cdot\Omega)\end{array}}{\Gamma \xrightarrow{\mathsf{FR}} [G_1 \land G_2]\cdot\Omega}\ {\land}R$$

Then we construct

$$\dfrac{\begin{array}{c}\mathsf{IH}_3(\mathcal{D}_1) \\ \Gamma \xrightarrow{\mathsf{FR}} [G_1]\end{array} \qquad \dfrac{\begin{array}{c}\mathsf{IH}_3(\mathcal{D}_1) \\ \Gamma \xrightarrow{\mathsf{FR}} [G_2 \land \bigwedge\Omega]\end{array}}{\Gamma \xrightarrow{\mathsf{FR}} [G_2]}\ \mathsf{INV}}{\Gamma \xrightarrow{\mathsf{FR}} [G_1 \land G_2]}\ {\land}R \qquad \text{and} \qquad \dfrac{\begin{array}{c}\mathsf{IH}_3(\mathcal{D}_1) \\ \Gamma \xrightarrow{\mathsf{FR}} [G_2 \land \bigwedge\Omega]\end{array}}{\Gamma \xrightarrow{\mathsf{FR}} [\bigwedge\Omega]}\ \mathsf{INV}$$

Here, INV refers to a use of inversion: if the premise is derivable then there is only a single rule that matches the conclusion (here ${\land}R$). In each case, we use one of the two premises that must have a derivation.

Also, because the induction hypothesis gives us two different derivations, it simultaneously justifies $\Gamma \xrightarrow{\mathsf{FR}} [G_1]$ and $\Gamma \xrightarrow{\mathsf{FR}} [G_2 \land \bigwedge\Omega]$.

Case (4):

$$\mathcal{D} = \quad \dfrac{\begin{array}{c}\mathcal{D}' \\ \Gamma \xrightarrow{\mathsf{FR}} [G]\cdot\Omega\end{array}}{\Gamma, [G \supset P] \xrightarrow{\mathsf{FL}} P\cdot\Omega}\ {\supset}L$$

Then we construct

$$\dfrac{\begin{array}{c}\mathsf{IH}_3(\mathcal{D}') \\ \Gamma \xrightarrow{\mathsf{FR}} [G]\end{array}}{\Gamma, [G \supset P] \xrightarrow{\mathsf{FL}} P}\ {\supset}L \qquad \text{and} \quad \begin{array}{c}\mathsf{IH}_3(\mathcal{D}') \\ \Gamma \xrightarrow{\mathsf{FR}} [\bigwedge\Omega]\end{array}$$

□

Corollary 2 *If* $\Gamma \xrightarrow{\mathsf{FR}} [G] \cdot \epsilon$ *then* $\Gamma \xrightarrow{\mathsf{FR}} [G]$.

Proof: From Theorem 1, part (3) for $\Omega = \epsilon$. □

4 Completeness of Goal Stacks

The proof in the other direction is often more complicated, or at least more difficult to arrive at the correct induction hypothesis. Again, we encourage you to try to come up with before looking at our proof.

The insight is the following. If we have, for example, $\Gamma \xrightarrow{\mathsf{FR}} [G]$ in the backward chaining system, then we should have $\Gamma \xrightarrow{\mathsf{FR}} [G]\cdot\epsilon$ in the goal stack system. This will however we insufficient to prove since goals stacks will accumulate during a derivation. So if $\Gamma \xrightarrow{\mathsf{FR}} [G]$, under which circumstances can we actually derive $\Gamma \xrightarrow{\mathsf{FR}} [G] \cdot \Omega$? We can do that if we also know $\Gamma \xrightarrow{\mathsf{CR}} \Omega$! That's because, intuitively. $\Gamma \xrightarrow{\mathsf{FR}} [G] \cdot \Omega$ will first derive G, bottom up, and then switch to deriving Ω.

Taking this into all three parts we get the following theorem. We can also see that it is (almost) the reverse of soundness.

Theorem 3 (Completeness of Goal Stacks)

(1) If $\Gamma \xrightarrow{\mathsf{FR}} [G]$ *and* $\Gamma \xrightarrow{\mathsf{CR}} \Omega$ *then* $\Gamma \xrightarrow{\mathsf{FR}} [G] \cdot \Omega$

(2) If $\Gamma, [D] \xrightarrow{\mathsf{FL}} P$ *and* $\Gamma \xrightarrow{\mathsf{CR}} \Omega$ *then* $\Gamma, [D] \xrightarrow{\mathsf{FL}} P \cdot \Omega$

(3) If $\Gamma \xrightarrow{\mathsf{C}} P$ *and* $\Gamma \xrightarrow{\mathsf{CR}} \Omega$ *then* $\Gamma \xrightarrow{\mathsf{CL}} P \cdot \Omega$

Proof: By simultaneous rule induction on the first given derivation $\mathcal{D}$ in parts (1)–(3) and arbitrary $\mathcal{E}$. We are given two derivations $\mathcal{D}$ (which varies from part to part) and $\mathcal{E}$ which derives $\Gamma \xrightarrow{\mathsf{R}} \Omega$. We show a few representative cases. The others are similar.

Case (1):

$$\mathcal{D} = \quad \dfrac{\begin{array}{cc} \mathcal{D}_1 & \mathcal{D}_2 \\ \Gamma \xrightarrow{\mathsf{FR}} [G_1] & \Gamma \xrightarrow{\mathsf{FR}} [G_2] \end{array}}{\Gamma \xrightarrow{\mathsf{FR}} [G_1 \wedge G_2]} \wedge R \qquad \text{and} \qquad \begin{array}{c} \mathcal{E} \\ \Gamma \xrightarrow{\mathsf{CR}} \Omega \end{array}$$

This is in some sense the trickiest case. We need to construct a derivation of $\Gamma \xrightarrow{\mathsf{FR}} [G_1 \wedge G_2]\cdot\Omega$ and therefore of $\Gamma \xrightarrow{\mathsf{FR}} [G_1]\cdot(G_2\cdot\Omega)$. But we cannot immediately apply the induction hypothesis on G_1, because we first need to get a derivation of $\Gamma \xrightarrow{\mathsf{CR}} G_2 \cdot \Omega$.

But we can get *this* from the induction hypothesis! So we construct:

$$\dfrac{\dfrac{\overset{\mathcal{D}_1}{\Gamma \xrightarrow{\mathsf{FR}} [G_1]} \qquad \dfrac{\dfrac{\overset{\mathcal{D}_2}{\Gamma \xrightarrow{\mathsf{FR}} [G_2]} \quad \overset{\mathcal{E}}{\Gamma \xrightarrow{\mathsf{CR}} \Omega}}{\Gamma \xrightarrow{\mathsf{FR}} [G_2]\cdot\Omega}\,\mathsf{IH}_1(\mathcal{D}_2,\mathcal{E})}{\Gamma \xrightarrow{\mathsf{CR}} G_2\cdot\Omega}\,FRCR}{\Gamma \xrightarrow{\mathsf{FR}} [G_1]\wedge(G_2\cdot\Omega)}\,\mathsf{IH}_1(\mathcal{D}_1,\_)}{\Gamma \xrightarrow{\mathsf{FR}} [G_1\wedge G_2]\cdot\Omega}\,\wedge R$$

Case (2):

$$\mathcal{D} = \quad \dfrac{\overset{\mathcal{D}'}{\Gamma \xrightarrow{\mathsf{FR}} [G]}}{\Gamma, [G \supset P] \xrightarrow{\mathsf{FL}} P}\,{\supset}L \qquad \text{and} \qquad \overset{\mathcal{E}}{\Gamma \xrightarrow{\mathsf{CR}} \Omega}$$

Then we construct

$$\dfrac{\overset{\mathsf{IH}_1(\mathcal{D}',\mathcal{E})}{\Gamma \xrightarrow{\mathsf{FR}} [G]\cdot\Omega}}{\Gamma, [G \supset P] \xrightarrow{\mathsf{FL}} P\cdot\Omega}\,{\supset}L$$

Case (2):

$$\mathcal{D} = \quad \dfrac{}{\Gamma, [P] \xrightarrow{\mathsf{FL}} P}\,\mathsf{id} \qquad \text{and} \qquad \overset{\mathcal{E}}{\Gamma \xrightarrow{\mathsf{CR}} \Omega}$$

Then we construct

$$\dfrac{\overset{\mathcal{E}}{\Gamma \xrightarrow{\mathsf{CR}} \Omega}}{\Gamma, [P] \xrightarrow{\mathsf{FL}} P\cdot\Omega}\,\mathsf{id}$$

□

5 Prolog Implementation

You can find the straightforward Prolog implementation of the goal stacks in Listing **??**. Code that also includes examples is in goalstack.pl.

```
mem(Dcopy, [D | Gamma]) :- copy_term(D, Dcopy).
mem(Dcopy, [_ | Gamma]) :- mem(Dcopy, Gamma).

unify(P,Q) :- unify_with_occurs_check(P,Q).
% unify(P,P).  % Prolog, but logically unsound

chooseR(Gamma, []).
chooseR(Gamma, [G|Omega]) :- focusR(Gamma, G, Omega).

focusR(Gamma, and(G1,G2), Omega) :- focusR(Gamma, G1, [G2|Omega]).
focusR(Gamma, atom(P), Omega) :- chooseL(Gamma, atom(P), Omega).

chooseL(Gamma, atom(P), Omega) :-
    mem(D, Gamma),
    focusL(Gamma, D, atom(P), Omega).

focusL(Gamma, atom(Q), atom(P), Omega) :-
    unify(Q,P),
    chooseR(Gamma, Omega).
focusL(Gamma, imp(G,atom(Q)), atom(P), Omega) :-
    unify(Q,P),
    focusR(Gamma, G, Omega).
```

Listing 1: Metainterpreter for Horn clauses using goal stacks