# Focusing and Polarity

We will step back from the computational interpretation and talk about the proof system today.

needed only for atomic formula A

$$\frac{}{\Gamma ; A \vdash A}\ id_A \qquad \frac{\Gamma ; \Delta \vdash A \quad \Gamma ; \Delta', A \vdash C}{\Gamma ; \Delta, \Delta' \vdash C}\ cut_A$$

$\boxed{\text{eliminable}}$

Recall that when we originally introduced these connectives, we tested them by checking if a cut of a more complicated expression could be simplified to the original cut. So a question you might you ask is, whether or not we need cut at all. The answer is you don't actually need this: if you have any proof, and that proof has (many) uses of cut, you can always reduce it to a proof without cut. Computationally, that says we can reduce communicating processes to "done" processes with no communication to do. (But note that when you compute, you don't reduce underneath, in the same way you don't reduce under a lambda in functional programming. Cut elimination says you can reduce everything to normal form, not just things at the top level.

From the proof search perspective, this is very powerful. In proof search, cut can be thought of having to "fabricate" an intermediate goal out of thin air. There is no way for proof search to figure this out! But since we know that we can get to the result without cuts, we know we don't have to make any lemmas.

Can we also eliminate identity? Well, in some cases you will need it; for example, if you have atomic formulas in the front and back--you cannot break it down further.

Gentzen 1935 proved cut-elimination (in fact, linear logic was hiding in his formalism, since he was explicit about using formulas twice. He had a strong and weak version of cut-elimination; the strong version was what we just looked about.

## So how do we proceed with proof search?

Focusing was invented by Andreoli 1991 in linear logic, an important development that didn't come from classical logic (however, we now know that this is widely applicable to many logics). It is often used as a test: if your rules do not admit focusing, it is questionable whether or not they are good rules.

Here is a statement, where we have multiple choices of proof, even without cut:

$$\dfrac{\dfrac{\dfrac{}{A, A \multimap B, B \multimap C, C \multimap D \vdash D}}{A \otimes (A \multimap B) \otimes (B \multimap C) \otimes (C \multimap D) \vdash D}}{\vdash A \multimap (A \multimap B) \otimes (B \multimap C) \otimes (C \multimap D) \multimap D}$$

three applications of left rule possible. all of them work

one step of tensor is forced

not much choice

With focusing, we will be able to reduce the number of possible proofs to one!

negative

$$\dfrac{\Delta, A \vdash B}{\Delta \vdash A \multimap B} \multimap R$$

Idea: we can always apply this rule. That is to say, when we see a linear implication on the rhs, ignore the context and just immediately apply it.

How might we prove this? Intuitively, we don't want to miss out on any possible proofs by prematurely committing to a course of action. Formally, we want to show a property that this rule is invertible, that is to say, we can take the inference both forwards and backwards.

$$\dfrac{\Delta \vdash A \multimap B \qquad \dfrac{\dfrac{}{A \vdash A}\, id_A \qquad \dfrac{}{B \vdash B}\, id_B}{A, A \multimap B \vdash B} \multimap L}{\Delta, A \vdash B} cut_{A \multimap B}$$

But notice there is a proof of this formula without cut, by application of cut elimination!

$\boxed{\text{Negative — Right Rule is invertible}}$

$\boxed{\text{Positive — Left Rule is invertible}}$

Q: Why do negative and positive cover all the cases?
A: I don't know! This has been mysterious to me ever since I started working on this. But it always seems to be true.

Q: Are there any connectives which are simultaneously positive and negative?
A: Yes. In linear logic, all connectives are uniquely negative/positive; in intuitionistic logic, conjunction is positive and negative, and you can decide how you want to use it. In classical logic, disjunction is also positive and negative. The less distinctions your logic makes, the more negative positive things are.
Q: Why is disjunction in classical logic both polarized?
A: In classical logic, you can hedge your bet, where you don't have to pick which conclusion you're going to prove immediately. To find rules which are positive and negative, you must go to quantifiers; all the other rules are agnostic.

$$\frac{\Gamma \vdash \Sigma, A, B}{\Gamma \vdash \Sigma, A \vee B}$$

positive

$$\frac{}{\cdot \vdash 1}\ 1R$$

1-right is not invertible, because although the rule per se is invertible, we actually need it to be invertible under any context.

$$\Delta \overset{?}{\vdash} 1 \quad \text{vs.} \quad \cdot \vdash 1$$

Q: Is there a way to tell what the polarity without looking at the rules?
A: No, not really. The meaning of the connectives are determined by the inference rules. So if I want to know if something is positive/negative, then I have to look at the rules.
Q: But isn't there a way of characterizing the quantifiers without rules?
A: Well, you might define two connectives with different rules, but they may be equivalent.

positive

$$\frac{\Delta \vdash A \qquad \Delta' \vdash B}{\Delta, \Delta' \vdash A \otimes B}\ \otimes R$$

Tensor-right is not invertible, there is an easy counter-example:

$$\frac{? \vdash A \qquad ? \vdash B}{\underbrace{A \otimes B}_{\text{hypothesis not splittable}} \vdash A \otimes B}$$

Essentially, all of the proofs and all of the counter-examples look the same.

negative

$$\frac{\Delta, A \vdash C}{\Delta, A \& B \vdash C}\ \& L_1 \qquad \frac{\Delta, B \vdash C}{\Delta, A \& B \vdash C}\ \& L_2$$

With-left clearly not invertible, since we're losing information.

You can do the rest of the proofs yourself, but to sum up:

| | |
|---|---|
| Negative | $A \multimap B,\ A \& B,\ \top,\ \forall x.A,\ P^-$ |
| Positive | $A \otimes B,\ 1,\ A \oplus B,\ \emptyset,\ \exists x.A,\ P^+,\ !A$ |

One of Andreoli's important insights was that atomic propositions could be either positive or negative: it is your choice!

There is something interesting going on with the exponential:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{A; A \vdash A}\ id_A}{A; \cdot \vdash A}\ copy}{A; \cdot \vdash !A}\ !R}{\cdot; !A \vdash !A}\ !L$$

there is a stutter step here, which doesn't show up in any of the other identity proofs. This will be important in more detailed accounts of polarizations, see Frank's lecture notes

Let's now see how focusing works, with the assumption that all of the propositions are negative:

$$[A] \vdash A$$
$$A \vdash A$$
$$A \vdash [A] \qquad [B] \vdash B$$
$$A, [A \multimap B] \vdash B$$
$$A, A \multimap B \vdash [B] \qquad [C] \vdash C$$
$$A, A \multimap B, [B \multimap C] \vdash C$$
$$A, A \multimap B, B \multimap C \vdash C \quad \text{blur}$$
$$A, A \multimap B, B \multimap C \vdash [C] \qquad [D] \vdash D$$
$$A, A \multimap B, B \multimap C, [C \multimap D] \vdash D \quad \text{focus}$$
$$A, A \multimap B, B \multimap C, C \multimap D \vdash D \quad \otimes R \times 3 \ \text{(invertible)}$$
$$A \otimes (A \multimap B) \otimes (B \multimap C) \otimes (C \multimap D) \vdash D \quad \multimap R \ \text{(invertible)}$$
$$\vdash A \otimes (A \multimap B) \otimes (B \multimap C) \otimes (C \multimap D) \multimap D$$

work on focused terms

Note that this doesn't discharge immediately; we have to move the focus over.

Now we have no more work to do, and the focus is on the rhs, so we can blur and pick a new focus.

After blasting through the invertible rules, we now have a choice of what to focus on the left hand side. Informally, we MUST pick the implication which has D in the goal.

This style of "goal-directed" search is at the heart of Prolog: we apply rules only when they allow us to derive what we are looking for in the goal.

If you want to represent a proof search strategy, you modify the rules so that you can only apply them if you are following the proof search strategy. So we need a focus rule, a blur rule, and modify some of our rules. So now we'll write down the rules of focusing.

$$\frac{}{[P^-] \vdash P}\ id^- \qquad\qquad \frac{}{P \vdash [P^+]}\ id^+$$

$$\frac{\Delta, [A^-] \vdash C}{\Delta, A^- \vdash C}\ focus_L \qquad\qquad \frac{\Delta, A^+ \vdash C}{\Delta, [A^+] \vdash C}\ blur_L$$

$$\frac{\Delta \vdash [C^+]}{\Delta \vdash C^+}\ focus_R \qquad\qquad \frac{\Delta \vdash C^-}{\Delta \vdash [C^-]}\ focus_L$$

We simply want to modify the non-invertible rules to only apply when a formula is in focus.

$$\frac{\Delta \vdash [A] \qquad \Delta', [B] \vdash C}{\Delta, [A \multimap B] \vdash C}\ \multimap L$$

see Frank's lecture notes for the rest!

The important part is we focus on one formula, and then we chain formulas on that focus.

There is an extremely strong restriction here, which is that this focusing strategy is complete. This is so powerful we can use our proof search language as a programming language. Returning to our previous example, suppose they are all positive:

$$A^+, B^+, C^+, D^+ \quad \text{positive}$$

Our previous proof would fail! But we would be able to focus on a different element (A -o B instead of C -o D), and the proof does go through. Notice that our hand is completely forced, since if we pick a bad focus, the proof will immediately fail. When all the atomic propositions are negative, you get Prolog (which is goal directed search); when all the atomic propositions are positive, you get Datalog (which is forward inference/chaining). (Well, there are some differences because Prolog/Datalog are not really linear. But there are linear versions which behave this way.

Note that we don't have to decide if the atoms are positive or negative uniformly; Andreoli showed we could do it any way we wanted to. We saw this in the first lecture, when we needed to define natural numbers.

$$\forall x.\ nat^-(x) \multimap nat^-(s(x))$$
$$nat^-(0)$$

If you consider everything positive, then you will generate all the natural numbers and never quiesce. But if you do goal directed search, nat will work fine. So nat should be negative.

This situation also showed up in our sorting algorithm:

$$\frac{elem^+(a,n,b) \quad elem^+(b,k,c) \quad gt^-(n,k)}{elem^+(a,k,b) \quad elem^+(b,n,c)}$$

Q: When you assign positive/negative, what are you actually doing?
A: Andreoli says each atomic proposition can be negative/positive. I don't actually know a good way of exploiting that; I only know how to do this for predicate. I do it per predicate; every result of a predicate is positive/negative. But hypothetically you could say one is negative, and two is positive. (That is to say, there doesn't seem to be a use for it.

I want to revisit concurrent functional programming for a moment. We just imposed a constraint where computation is proof search, a search so restricted that it can be interpeted as a program. This has a lot of concurrency, and we used this when we used logic programming to explain functional programming. Let's do the concurrent part: what does focusing/polarity tell us about the functional interpretation of proofs and connectives?

For inversion, the information content of applying an invertible rule is zero. You don't gain anything, you don't lose anything. This means that in a computational interpretation, no information is implicit in the rule. So it couldn't be an output; the right rule for linear implication would need some information, but the rule doesn't have any. So it must be an INPUT, they must receive information.

| | |
|---|---|
| $A \multimap B$ | input of $A$ |
| $A \& B$ | input of $\pi_1/\pi_2$ |
| $\forall x{:}\tau.A$ | input of $x$ |

Positive connectives make a choice, they produce information, that choice is the output of the connective.

| | |
|---|---|
| $A \otimes B$ | output of partition of resources |
| $A \oplus B$ | output of which one I'm supplying |
| $\exists x{:}\tau.A$ | output of the witness |

This closes the circle back to what Bob Harper talked about, with regards to positive and negative connectives.

Q: If you have proof terms in a system with focusing, is there a constructor that corresponds to focusing/blur?
A: If you look at just the negative fragment, then what happens is when you focus on something negative (with assumption x : A), you get x at the head of the term, and each application of implication left adds an argument to it.

the function $\longrightarrow (x \cdot (M ; N)) : C$

pf of $A$ ↗ ↖ pf of $B$

$x : A \multimap B \multimap C$

$((x \underset{A}{M}) \underset{B}{N}) : C$

in natural deduction

bijection when only considering negatives

Q: What is the relationship of this and CBN and CBV strategies?
A: Yes, there is. Unfortunately, I cannot explain it in two minutes.

Q: In classic linear logic, there's another connective for disjunction. Why?
A: We cannot write down the left and right rules, when there is a singleton on the right hand side. However, in classical linear logic, you cannot have the linear implication right rule, since you need to force the right-hand to be singleton. As usual, you can express classical LL in intuitionistic language using continuation passing style (it shows up as tensor on the other side); the reverse is not true. I don't really know what "why not" is supposed to mean, maybe it's "critical regions", but I don't really know. I can write down rules for executing it, but I can't read anything out of it.

Q: If you have inductive types, what are their polarity?
A: This goes back to a question from the first lecture, which is when you string together multiple connectives, what is their polarity?

$\mu\alpha.\, 1 \oplus \alpha$ [Baelde09]

Are these inference rules consistent? Yes: if you have this formula, you can derive this rule. You also need to consider the opposite; if you have focusing, you can do this. So a theorem states you can always go back and forth between a rule and the formula, so long as the formula satisfies a condition: that is it is a bipole, there is at most one change of polarity.

$$\frac{A \qquad B}{C \qquad D}$$

$$\underbrace{!(A \otimes B \multimap C \otimes D)}_{\text{bipole}}$$

but with some waffling about what to do about the exponentials. I'd like to say this is just a persistent rule, but we haven't written those rules. Bipoles are an Andreoli 2001 paper.