Lecture Notes on Negative Types

15-814: Types and Programming Languages Frank Pfenning

> Lecture 19 Thursday, November 11, 2021

1 Introduction

We continue the investigation of shared memory concurrency by adding *negative types*. In our language so far they are functions $\tau \rightarrow \sigma$, lazy pairs $\tau \otimes \sigma$, and universal types $\forall \alpha. \tau$.

2 Review of Positives

We review the types so far, with a twist: we annotate every address that we write to with a superscript^W and every address we read from with a superscript^R.

The configurations are unordered and we think of "," as an associative and commutative operator with unit ".". Since we have changed our notation, we summarize the translation and the transition rules.

LECTURE NOTES

$$\begin{split} \llbracket x \rrbracket d &= d^{W} \leftarrow x^{R} \\ \llbracket \langle \rangle \rrbracket d &= d^{W} \cdot \langle \rangle \\ \llbracket \text{(case } e \ (\langle \rangle \Rightarrow e') \rrbracket d = x \leftarrow \llbracket e \rrbracket x \ ; \\ \text{ case } x^{R} \ (\langle \rangle \Rightarrow \llbracket e' \rrbracket d) \\ \llbracket \langle e_{1}, e_{2} \rangle \rrbracket d &= x_{1} \leftarrow \llbracket e_{1} \rrbracket x_{1} \ ; \\ x_{2} \leftarrow \llbracket e_{1} \rrbracket x_{2} \ ; \\ d^{W} \cdot \langle x_{1}, x_{2} \rangle \\ \llbracket \text{(case } e \ (\langle x_{1}, x_{2} \rangle \Rightarrow e') \rrbracket d = x \leftarrow \llbracket e \rrbracket x \ ; \\ \text{ case } x^{R} \ (\langle x_{1}, x_{2} \rangle \Rightarrow \llbracket e' \rrbracket d) \\ \llbracket j \cdot e \rrbracket d = x \leftarrow \llbracket e \rrbracket x \ ; \\ d^{W} \cdot (j \cdot x) \\ \llbracket \text{(case } e \ (i \cdot x \Rightarrow e_{i})_{i \in I} \rrbracket d = x \leftarrow \llbracket e \rrbracket x \ ; \\ \text{ case } x^{R} \ (i \cdot x \Rightarrow \llbracket e_{i} \rrbracket d)_{i \in I} \\ \llbracket \text{fold } e \rrbracket d = x \leftarrow \llbracket e \rrbracket x \ ; \\ d^{W} \cdot (\text{fold } x) \\ \llbracket \text{(case } e \ (\text{fold } y \Rightarrow e') \rrbracket d = x \leftarrow \llbracket e \rrbracket x \ ; \\ \text{ case } x^{R} \ (\text{fold } y \Rightarrow \llbracket e' \rrbracket d) \\ \llbracket \text{(case } e \ (\text{fold } y \Rightarrow \llbracket e') \rrbracket d = x \leftarrow \llbracket e \rrbracket x \ ; \\ \text{ case } x^{R} \ (\text{fold } y \Rightarrow \llbracket e' \rrbracket d) \end{split}$$

We only have four transition rules for configurations, in addition to explaining how values are matched against continuations.

 $\begin{array}{rcl} & \operatorname{proc} d \; (x \leftarrow P \; ; \; Q) & \mapsto & \operatorname{proc} c \; ([c/x]P), \operatorname{proc} d \; ([c/x]Q) & (c \; \operatorname{fresh}) \\ \operatorname{!cell} c \; V, & \operatorname{proc} d \; (d^W \leftarrow c^R) & \mapsto & \operatorname{!cell} d \; V \\ & \operatorname{proc} d \; (d^W.V) & \mapsto & \operatorname{!cell} d \; V \\ \operatorname{!cell} c \; V, & \operatorname{proc} d \; (\operatorname{case} c^R \; K) & \mapsto & \operatorname{proc} d \; (V \triangleright K) \end{array}$

3 Functions

As the first negative type we consider function $\tau \rightarrow \sigma$. How do we translate an abstraction $\lambda x. e$? The translation must actually take *two* arguments: one is the original argument *x*, the other is the destination where the result of

LECTURE NOTES

the functional call should be written to. And the process $[\lambda x. e] d$ must write the translation of the function to destination *d*.

Before we settle on the syntax for this, consider how to translate function application.

 $\llbracket e_1 e_2 \rrbracket d = x_1 \leftarrow \llbracket e_1 \rrbracket x_1 ;$ $x_2 \leftarrow \llbracket e_2 \rrbracket x_2 ;$

How should we complete this translation?

We know that after $\llbracket e_1 \rrbracket x_1$ has completed the cell x_1 will contain *a function of two arguments*. The *first* argument is the original argument, which we find in x_2 after $\llbracket e_2 \rrbracket x_2$ has completed. The *second* argument is the destination for the result of the function application, which is *d*. So we get:

 $\begin{bmatrix} e_1 e_2 \end{bmatrix} d = x_1 \leftarrow \begin{bmatrix} e_1 \end{bmatrix} x_1 ; \\ x_2 \leftarrow \begin{bmatrix} e_2 \end{bmatrix} x_2 ; \\ x_1^R . \langle x_2, d \rangle$

This looks just like eager pairs, except that we *read* from x_1 instead of writing to it. To retain the analogy, we write the translation of a function using case, but *writing* the (single) branch of the case expression to memory.

 $\llbracket \lambda x. e \rrbracket d = \mathsf{case} \; d^W \; (\langle x, y \rangle \Rightarrow \llbracket e \rrbracket y)$

The transition rules for these new constructs just formalize the explanation.

$$\begin{array}{l} \operatorname{proc} d \; (\operatorname{case} \; d^W \; (\langle x, y \rangle \Rightarrow P)) \mapsto \operatorname{!cell} d \; (\langle x, y \rangle \Rightarrow P) & (\rightarrow R) \\ \operatorname{!cell} c \; (\langle x, y \rangle \Rightarrow P), \operatorname{proc} d \; (c^R. \langle c_1, d \rangle) \mapsto \operatorname{proc} d \; ([c_1/x, d/y]P) & (\rightarrow L^0) \end{array}$$

As an example, we consider the expression $(\lambda x. x) \langle \rangle$.

$$\begin{split} \llbracket (\lambda x. x) \left\langle \right\rangle \rrbracket d_0 &= x_1 \leftarrow \llbracket \lambda x. x \rrbracket x_1 ; \\ & x_2 \leftarrow \llbracket \left\langle \right\rangle \rrbracket x_2 ; \\ & x_1^R. \langle x_2, d_0 \rangle \\ &= x_1 \leftarrow \operatorname{case} x_1^W \left(\langle x, y \rangle \Rightarrow \llbracket x \rrbracket y \right) ; \\ & x_2 \leftarrow x_2^W. \langle \right\rangle ; \\ & x_1^R. \langle x_2, d_0 \rangle \\ &= x_1 \leftarrow \operatorname{case} x_1^W \left(\langle x, y \rangle \Rightarrow y^W \leftarrow x^R \right) \\ & x_2 \leftarrow x_2^W. \langle \right\rangle ; \\ & x_1^R. \langle x_2, d_0 \rangle \end{split}$$

LECTURE NOTES

THURSDAY, NOVEMBER 11, 2021

;

Let's execute the final process from with the initial destination d_0 .

 $\begin{array}{ll} \operatorname{proc} d_0 \left(x_1 \leftarrow \operatorname{case} x_1^W \left(\ldots \right) ; x_2 \leftarrow x_2^W . \left\langle \right\rangle ; \ldots \right) \\ \mapsto & \operatorname{proc} d_1 \left(\operatorname{case} d_1^W \left(\left\langle x, y \right\rangle \Rightarrow y^W \leftarrow x^R \right) \right), \\ & \operatorname{proc} d_0 \left(x_2 \leftarrow x_2^W . \left\langle \right\rangle ; d_1^R . \left\langle x_2, d_0 \right\rangle \right) \\ \mapsto^2 & \operatorname{lcell} d_1 \left(\left\langle x, y \right\rangle \Rightarrow y^W \leftarrow x^R \right), \\ & \operatorname{proc} d_2 \left(d_2^W . \left\langle \right\rangle \right), \\ & \operatorname{proc} d_0 \left(d_1^R . \left\langle d_2, d_0 \right\rangle \right) \\ \mapsto & \operatorname{lcell} d_1 \left(\left\langle x, y \right\rangle \Rightarrow y^W \leftarrow x^R \right), \\ & \operatorname{lcell} d_2 \left\langle \right\rangle, \\ & \operatorname{proc} d_0 \left(d_0^W \leftarrow d_2^R \right) & (\operatorname{from} \left[d_2 / x, d_0 / y \right] (y^W \leftarrow x^R)) \\ \mapsto & \operatorname{lcell} d_1 \left(\left\langle x, y \right\rangle \Rightarrow y^W \leftarrow x^R \right), \\ & \operatorname{lcell} d_2 \left\langle \right\rangle \\ & \operatorname{lcell} d_2 \left\langle \right\rangle \\ & \operatorname{lcell} d_0 \left\langle \right\rangle \end{array}$

In the final configuration we have cell d_0 holding the final result $\langle \rangle$, which is indeed the result of evaluating $(\lambda x. x) \langle \rangle$. We also have some newly allocated intermediate destinations d_1 and d_2 that are preserved, but could be garbage collected if we only retain the cells that are reachable from the initial destination d_0 which now holds the final value.

4 Store Revisited

In our table of process expression, two things stand out. One is that functions are exactly like pairs, except that the role of reads and writes are reversed. The other is that a cell may now contain something of the form $(\langle y, z \rangle \Rightarrow P)$.

Processes	P	::=	$x \leftarrow P; Q$		allocate/spawn
			$x^W \leftarrow y^R$		сору
			$x^W.\langle \rangle$	\mid case x^{R} $(\langle \rangle \Rightarrow P)$	(1)
			$x^W.\langle y,z angle$	\mid case $x^{R}\left(\left\langle y,z ight angle \Rightarrow P ight)$	(\times)
			$x^W.(j \cdot y)$	case $x^R \ (i \cdot y \Rightarrow P_i)_{i \in I}$	(+)
			$x^W.fold(y)$	$ \operatorname{case} x^R \left(\operatorname{fold}(y) \Rightarrow P \right)$	(μ)
			$x^R.\langle y,z\rangle$	$\mid case\; x^W\; (\langle y,z\rangle \Rightarrow P)$	(\rightarrow)

We can refactor this into a more uniform presentation, even though not all of the syntactically legal forms have corresponding types in the current

LECTURE NOTES

language.

Small values
$$V$$
 $::=$ $\langle \rangle \mid \langle a_1, a_2 \rangle \mid i \cdot a \mid \text{fold } a$ Continuations K $::=$ $(\langle \rangle \Rightarrow P) \mid (\langle x_1, x_2 \rangle \Rightarrow P) \mid (i \cdot x_i \Rightarrow P_i)_{i \in I} \mid (\text{fold } x \Rightarrow P)$ Cell contents W $::=$ $V \mid K$ Configurations C $::=$ $\cdot \mid C_1, C_2 \mid \text{proc } d \mid P \mid \text{cell } c \mid W$

There is now a legitimate concern that the contents of cells in memory is no longer "small", because a program *P* could be of arbitrary size. At a lower level of abstraction, continuations would probably be implemented as *closures*, that is, a pairs consisting of an environment and the address of code to be executed. The translation to get us to this form is called *closure conversion*, which we might discuss in a future lecture. For now, we are content with the observation that, yes, we are violating a basic principle of fixed-size storage and that it can be mitigated (but is not completely solved) through the introduction of closures.

In our example of $(\lambda x. x) \langle \rangle$ the continuation has the form $(\langle x, y \rangle \Rightarrow y^W \leftarrow x^R)$ which is a closed process. This can be directly compiled to a function that takes two addresses x and y and writes the contents of x into y. So at least in this special case the contents of the cell d_1 could simply be the address of this piece of code.

The symmetry between eager pairs (positive) and functions (negative) stems from the property that in logic we have $A \vdash B \supset C$ if and only if $A \times B \vdash C$ (where \times is a particular form of conjunction). Or, we can chalk it up to the isomorphism $\tau \rightarrow (\sigma \rightarrow \rho) \cong (\tau \times \sigma) \rightarrow \rho$: an arrow on the right behaves like a product on the left.

One can ask if similarly symmetric constructors exists for 1, +, and μ and the answer is yes. It turns out that lazy records are dual to sums and there is a type \perp that is dual to 1 (see Exercises 1 and 2). There may even be a lazy analogue of recursive types that exhibits the same kind of symmetry and maybe useful to model so-called corecursive types (see Exercise 3).

We postpone discussion on the typing of process expression, cells, and configurations until the next lecture when we consider analogues of the

LECTURE NOTES

progress and preservation theorems.

5 Typing

Before writing an example, it may be helpful to introduce typing for configuration. The judgment has the form

$$\Psi \vdash \mathcal{C} :: \Delta$$

where the processes and cells in C may read from Ψ and either define (cells) or write to (processes) Δ . Both of these are contexts give *types* to *addresses*:

$$\Psi ::= a : \tau \mid \cdot \mid \Psi_1, \Psi_2$$

We start with the rule for composition of configurations. Even though configurations, as written, are *unordered*, their typing derivation will impose an order in that a process writing to an address comes before (to the left) or a process reading from an address. Cells are ordered correspondingly: lcell a W comes before any reader of a, and any cell referenced by W must come before it.

$$\frac{\Psi_0 \vdash \mathcal{C}_1 :: \Psi_1 \quad \Psi_1 \vdash \mathcal{C}_2 :: \Psi_2}{\Psi_0 \vdash \mathcal{C}_1, \mathcal{C}_2 :: \Psi_2} \text{ tp/join } \qquad \frac{\Psi \vdash (\cdot) :: \Psi}{\Psi \vdash (\cdot) :: \Psi} \text{ tp/empty}$$

Because cells might have multiple readers, in the typing of processes and cells we have to propagate all channels from the left of the judgment to the right.

$$\frac{\Psi \vdash P :: (d : \tau)}{\Psi \vdash \operatorname{proc} d P :: (\Psi, d : \tau)} \operatorname{tp/proc}$$

For cells, we distinguish the cases of a small value V or a continuation K. In either case, to avoid redundancy between various rules, we just reconstitute a process that would create such a cell and type that.

$$\frac{\Psi \vdash c^W.V :: (c:\tau)}{\Psi \vdash !\mathsf{cell} \ c \ V :: (\Psi, c:\tau)} \ \mathsf{tp/cell/val} \qquad \frac{\Psi \vdash \mathsf{case} \ c^W \ K :: (c:\tau)}{\Psi \vdash !\mathsf{cell} \ c \ K :: (\Psi, c:\tau)} \ \mathsf{tp/cell/cont}$$

We already discussed the typing rules for the positives in the last lecture, so it remains to show typing for the negatives. The only one so far is functions;

LECTURE NOTES

others are discussed in the exercises.

$$\begin{split} & \frac{\Gamma, y: \tau \vdash P :: (z:\sigma)}{\Gamma \vdash \mathsf{case} \; x^W \; (\langle y, z \rangle \Rightarrow P) :: (x:\tau \to \sigma)} \to R \\ & \frac{x: \tau \to \sigma \in \Gamma \quad y: \tau \in \Gamma}{\Gamma \vdash x^R . \langle y, z \rangle :: (z:\sigma)} \to L^0 \end{split}$$

Under the Curry-Howard interpretation into the semi-axiomatic sequent calculus, this corresponds to the following two rules

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \supset B} \supset R \qquad \qquad \frac{\Gamma, A \supset B, A \vdash B}{\Gamma, A \supset B, A \vdash B} \supset L^0$$

6 Preservation and Progress

We see the rules are arranged so that $\Psi \vdash C :: \Delta$ implies that $\Psi \subseteq \Delta$. In the preservation theorem we need to account for the possibility that a new cell is allocated which would then appear in Δ' with its type but not in Δ .

While configurations are not explicitly ordered, a typing derivation imposes some ordering constraints. In particular, a cell (or the writer of a cell), always precedes a reader of a cell in the left-to-right order of the typing derivation.x

In this lecture we only state progress and preservation; we may come back later to prove them when our language is complete.

Theorem 1 (Preservation)

If $\Gamma \vdash \mathcal{C} :: \Delta$ and $\mathcal{C} \mapsto \mathcal{C}'$ then $\Gamma \vdash \mathcal{C} :: \Delta'$ for some $\Delta' \supseteq \Delta$.

To state progress, we should reflect on what plays the role of a *value* in our usual formulation of progress. But it turns out to be easy: it is a configuration consisting entirely of cells and no processes. We call such a configuration *final*. Clearly, such a configuration cannot take a step. The usual notion of a *closed expression* that we start with is replace by a configuration that does not rely (that is, may read from) any external addresses.

Theorem 2 (Progress) *If* $\cdot \vdash C :: \Delta$ *then either* $C \mapsto C'$ *or* C *is final.*

LECTURE NOTES

7 Example: A Pipeline

As a simple example for concurrency in this language we consider setting up a (very small) pipeline. We consider a sequence of bits

 $bits = \mu \alpha. (\mathbf{b0} : \alpha) + (\mathbf{b1} : \alpha) + (\mathbf{e} : 1)$

(which also happen to be isomorphic to binary numbers). We assume there is a process $flip : bits \rightarrow bits$ that just flips every bit. We show this below. For now, our goal is to understand how to compose two such processes in a pipeline.

Assume there is a cell

!cell *flip*
$$K_{flip}$$
 : *bits* \rightarrow *bits*

This means that $K_{flip} = (\langle x, y \rangle \Rightarrow P)$ where x : bits is address of the argument and y : bits is the destination for the result.

Then we can compose two of these as

$$\begin{split} K_{\textit{flip2}} = \langle x, z \rangle \Rightarrow \\ y \leftarrow \textit{flip}^R. \langle x, y \rangle \\ \textit{flip}^R. \langle y, z \rangle \end{split}$$

In the picture below we see the two *flip* processes running, after the code for *flip2* has executed but neither of these has taken any action yet. The process on the left reads from x and writes to y while the process on the right reads from y and writes to z. The destinations y and z have been allocated but have not yet been written to. Cell x contains the sample input, which is the

LECTURE NOTES

memory representation of fold $(\mathbf{b0} \cdot \mathsf{fold} \ (\mathbf{e} \ \cdot \langle \rangle))$.

The left process now reads along x and allocates and writes along y. After it runs for a few steps, we might reach the following situation:



The green part here is the new part compared to the previous configuration.

LECTURE NOTES

It should be clear how each of the two processes translates into a proc object, while each filled cell corresponds to a cell object. The empty cells are addresses that have been allocated, but not yet written to, so they are not explicit in the configuration.

The two processes also run in parallel, which is how they form a pipeline. For example, after a few more steps we might reach the configuration (with the purple part being new):



The right process here lags behind the left one, which is possible since the semantics here is not synchronous. A cell can be read as soon as it is filled, but it may not be read immediately while other computations take place.

If we knew that the left process was the only reader along x (and any cells reachable from it) we could "garbage-collect" the cells that are no longer accessible and the situation would look as follows (assuming here

LECTURE NOTES



some process not shown could read the output z).

Instead of translating an expression, we write a *flip* process directly. For this purpose we have to decide how to handle recursion. There seem to be two solutions:

- 1. We add a process fix f. P which transitions to [fix f. P/f]P. This is entirely straightforward but requires process substitution in the dynamics.
- 2. We allow recursively defined processes

!cell flip K_{flip}

where K_{flip} refers back to its own cell with address *flip* to encode a recursive call.

We choose the latter option for this example, for variety, even though it would require more complicated typing rules for configurations.

It remains to define K_{flip} .



LECTURE NOTES

In the first branch, we have to allocate a fresh cell y' for the output and make a recursive call to fill it. We can also write $b1 \cdot y'$ to y.

$$\begin{split} K_{flip} &= \langle x, y \rangle \Rightarrow \mathsf{case} \ x \ (\ \mathbf{b0} \cdot x' \Rightarrow y' \leftarrow flip^R . \langle x', y' \rangle \ ; \\ & y^W . (\mathbf{b1} \cdot y') \\ & | \ \mathbf{b1} \cdot x' \Rightarrow \boxed{} \\ & | \ \mathbf{e} \cdot u \Rightarrow \boxed{} \end{split}$$

The branch for **b1** is symmetric to the first one.

$$\begin{split} K_{flip} &= \langle x, y \rangle \Rightarrow \mathsf{case} \; x \; (\; \mathbf{b0} \cdot x' \Rightarrow y' \leftarrow flip^R. \langle x', y' \rangle \; ; \\ & y^W. (\mathbf{b1} \cdot y') \\ & | \; \mathbf{b1} \cdot x' \Rightarrow y' \leftarrow flip^R. \langle x', y' \rangle \; ; \\ & y^W. (\mathbf{b0} \cdot y') \\ & | \; \mathbf{e} \cdot u \Rightarrow \boxed{} \end{split}$$

In the last case, we just write $\mathbf{e} \cdot u$ to y.

$$\begin{split} K_{flip} &= \langle x, y \rangle \Rightarrow \mathsf{case} \; x \; (\; \mathbf{b0} \cdot x' \Rightarrow y' \leftarrow flip^R. \langle x', y' \rangle \; ; \\ & y^W. (\mathbf{b1} \cdot y') \\ & | \; \mathbf{b1} \cdot x' \Rightarrow y' \leftarrow flip^R. \langle x', y' \rangle \; ; \\ & y^W. (\mathbf{b0} \cdot y') \\ & | \; \mathbf{e} \cdot u \Rightarrow \quad y^W. (\mathbf{e} \cdot u) \;) \end{split}$$

As shown in the previous section, we can compose two *flip* processes into a pipeline as follows:

$$K_{flip2} = \langle x, z \rangle \Rightarrow$$

$$y \leftarrow flip^{R} \langle x, y \rangle$$

$$flip^{R} \langle y, z \rangle$$

You may look back at the diagrams to visualize how the two processes work together, effectively communicating via the shared location y, which then becomes y', y'', etc. as the computation progresses and recursive calls are mad in both of them.

Under a sequential interpretation, where $x \leftarrow P$; Q waits until P has written to destination x before Q starts executing, all recursive calls in *flip* would have to be finished before the first bit of output is written. When we compose two, the inner one has to finish entirely, writing out the whole sequence of bits before the outer one can start. This is the behavior of the functional λx . *flip* (*flip* x) where the intermediate destination y and the final destination z remain unnamed.

LECTURE NOTES

Exercises

Exercise 1 For lazy records (as a generalization of lazy pairs) we introduce the following syntax in our language of expressions:

Types ::= $\dots | \&_{i \in I}(i : \tau_i)$ Expressions ::= $\dots | \langle i \Rightarrow e_i \rangle_{i \in I} | e \cdot j$

- 1. Give the typing rules and the dynamics (stepping rules) for the new constructs.
- 2. Extend the translation $\llbracket e \rrbracket d$ to encompass the new constructs. Your process syntax should expose the duality between eager sums and lazy records.
- 3. Extend the transition rules of the store-based dynamics to the new constructs. The translated form may permit more parallelism than the original expression evaluation, but when scheduled sequentially they should have the same behavior (which you do not need to prove).
- 4. Show the typing rules for the new process constructs.

Exercise 2 Explore what the rules and meaning of \perp as the formal dual of 1 in the process language should be, including whichever of the following you find make sense. If something does not make sense somehow, please explain.

- 1. Write out the new forms of process expressions.
- 2. Provide the store-based dynamics for the new process expressions.
- 3. Show the typing rules for the new process expressions.
- 4. Reverse-engineer new functional expressions in our original language so they translate to your new process expression. Show the rules for typing and stepping the new constructs.
- 5. Summarize and discuss what you found.

Exercise 3 In our expression language the fold *e* constructor for elements of recursive type is eager. Explore a new *lazy* roll *e* constructor which has type $\nu\alpha$. τ , providing:

1. Typing rules for roll and a corresponding destructor (presumably an unroll or case construct).

LECTURE NOTES

- L19.14
 - 2. Stepping rules for the new forms of expressions.
 - 3. A translation from the new forms of expressions to processes, extending the language of processes as needed
 - 4. Typing rules for the new forms of processes.
 - 5. Transition rules for the new forms of processes.