Day 5

1. Introducing Names

Let's talk about names.

$$\begin{aligned}
&\mathcal{X} \ni x \\
&\mathcal{V} \in v ::= z \\
&\mathcal{T} \ni t ::= z \mid t_1 + t_2 \mid t_1 \times t_2 \mid t_1 \div t_2 \mid x \mid \texttt{let } x = t_1 \texttt{ in } t_2
\end{aligned}$$

As before, we want:

- An evaluation relation
- An approximation of the evaluation relation that guarantees safety.

What are the problems?

- $\dfrac{}{x \Downarrow ??}$
- $\dfrac{t_1 \Downarrow v_1 \quad t_2 \Downarrow v_2}{\texttt{let } x = t_1 \texttt{ in } t_2 \Downarrow v_2}$ ... but where did v_1 go?

2. Substitution

First approach: *substitute* values into terms.

We define the substitution of a value v for a variable x in a term t (notation $t[v/x]$) as follows:

$$\begin{aligned}
y[v/x] &= \begin{cases} v & \text{if } x = y \\ y & \text{otherwise} \end{cases} \\
(t_1 \odot t_2)[v/x] &= t_1[v/x] \odot t_2[v/x] \qquad\qquad \odot \in \{+, \times, \div\} \\
(\texttt{let } y = t_1 \texttt{ in } t_2)[v/x] &= \begin{cases} \texttt{let } y = t_1[v/x] \texttt{ in } t_2 & \text{if } x = y \\ \texttt{let } y = t_1[v/x] \texttt{ in } [v/x]t_2 & \text{otherwise} \end{cases}
\end{aligned}$$

Relevant points:

- Relying on the inclusion of values in terms $\mathcal{V} \subseteq \mathcal{T}$. Could introduce explicit notation for this, but not even I am that pedantic.
- Shadowing of variables in `let`. (Intuition: bound names don't matter. Will pay off momentarily.)

Now, we are equipped to give our first meaning of variables and `let`:

$$\frac{t_1 \Downarrow v_1 \quad t_2[v_1/x] \Downarrow v_2}{\texttt{let } x = t_1 \texttt{ in } t_2 \Downarrow v_2}$$

- Substitution is a *meta-theoretic* notion: we don't have separate evaluation rules for $x[4/x]$ and 4, we treat those as the same term.

No rule for variables:

$$\frac{\dfrac{}{4 \Downarrow 4} \quad \dfrac{\dfrac{}{4 \Downarrow 4} \quad \dfrac{}{4 \Downarrow 4}}{4 \div 4 \Downarrow 1}}{\texttt{let } x = 4 \texttt{ in } x \div x \Downarrow 1}$$

So variables are always stuck terms: no derivation for $\texttt{let } x = 5 \texttt{ in } y \Downarrow z$ for any z.

3. α-Equivalence

Intuition: changing the names of local variables doesn't matter. Now, we're in a position to capture this idea formally.

We define *α-equivalence*—i.e., equivalence up to renaming of variables—by:

$$\frac{}{x \equiv_\alpha x} \qquad \frac{}{z \equiv_\alpha z} \qquad \frac{t_1 \equiv_\alpha t_1' \quad t_2 \equiv_\alpha t_2'}{t_1 \odot t_2 \equiv_\alpha t_1' \odot t_2'} \ (\odot \in \{+, \times, \div\})$$

$$\frac{t_1 \equiv_\alpha t_1' \quad t_2[z/x] \equiv_\alpha t_2'[z/y]}{\texttt{let } x = t_1 \texttt{ in } t_2 \equiv_\alpha \texttt{let } y = t_1' \texttt{ in } t_2'} \ (z \notin fv(t_1) \cup fv(t_2))$$

where the *free variables* of a term are intuitively those variables in the term not defined by an enclosing `let` statement:

$$fv(x) = \{x\} \qquad fv(t_1 \odot t_2) = fv(t_1) \cup fv(t_2), \quad \odot \in \{+, \times, \div\}$$
$$fv(z) = \emptyset \qquad fv(\texttt{let } x = t_1 \texttt{ in } t_2) = fv(t_1) \cup (fv(t_2) \setminus \{x\})$$

Why do we need a new (also called "fresh") variable in the `let` case? Mostly to avoid the possibility that x is already used in $t2'$.

Now we can make formal our intuition about α-equivalence:

Theorem. *If $t \equiv_\alpha t'$ and $t \Downarrow v$ then $t' \Downarrow v$.*

Proof. By structural induction on the derivation of $t \equiv_\alpha t'$:

- Case $\dfrac{}{x \equiv_\alpha x}$: the second hypothesis ($x \Downarrow v$) is impossible.
- Case $\dfrac{}{z \equiv_\alpha z}$: by definition of $\Downarrow$.
- Case $\dfrac{t_1 \equiv_\alpha t_1' \quad t_2 \equiv_\alpha t_2'}{t_1 \odot t_1' \equiv_\alpha t_2 \odot t_2'}$: If $t \Downarrow v$, then we have that $t_1 \Downarrow v_1$, $t_2 \Downarrow v_2$, and (abusing notation slightly) $v = v_1 \odot v_2$. Now, by the induction hypothesis, $t_1' \Downarrow v_1$, $t_2' \Downarrow v_2$, and finally by the definition of $\Downarrow$ we have $t' \Downarrow v$.

- Case $\dfrac{t_1 \equiv_\alpha t_1' \quad [z/x]t_2 \equiv_\alpha [z/y]t_2'}{\texttt{let } x = t_1 \texttt{ in } t_2 \equiv_\alpha \texttt{let } y = t_1' \texttt{ in } t_2'}$: By the induction hypothesis applied to the first subderivation we have $t_1 \Downarrow v_1$, $t_1' \Downarrow v_1$. Similarly, by the IH applied to the second subderivation, we have $t_2[z/x][v_1/z] \Downarrow v_2$ and $t_2'[z/y][v_1/z] \Downarrow v_2$. But the latter two expressions are equivalent (by tedious lemma) to $t_2[v_1/x]$ and $t_2'[v_1/y]$, so we have that the original terms evaluate to v_2 as well. □