July 29, 2016

1 Effects

R is a fixed set of resources. We define a resource as a language primitive with the authority to perform I/O operations. Π is the set of I/O operations. In this document we cannot dynamically create resources or operations on resources.

Members of R are denoted r; members of Π are denoted π. $r.\pi$ is syntactic sugar for (r, π) (for example, `FileIO.append` instead of $(\texttt{FileIO}, \texttt{append})$). An effect is a member of the pairs $R \times \Pi$. A set of effects is denoted by ε.

We say a piece of code C has the runtime effect $r.\pi$ if $r.\pi$ is called during the execution of C. C captures $r.\pi$ if it has the authority to call $r.\pi$ at some point during its execution. The static effects of C is an approximation of the runtime effects by our typing system. Later on we show the static effects of C give a conservative upper-bound on the runtime effects.

Types in our system are either resources or structural. Structural types are distinguished by what method declarations they have. The type with no methods is called `Unit`; its unique instance of denoted `unit`. Although they look similar in form, operations and methods are distinct. Methods can only be invoked by objects; operations can only be invoked by resources.

We make some simplifying assumptions about methods and operations. Methods take exactly one argument. If the argument is not specified it is assumed to be `unit`. Invoking some operation $r.\pi$ returns $\varnothing$. We don't model arguments to operations, so all operations are null-ary.

2 Static Semantics For Fully-Annotated Programs

In this first system every method in the program is explicitly annotated with its set of effects.

2.1 Grammar

$$
\begin{array}{lll}
e & ::= & x \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \textit{expressions} \\
& | & r \\
& | & \texttt{new}_{\sigma}\ x \Rightarrow \overline{\sigma = e} \\
& | & e.m(e) \\
& | & e.\pi \\
\\
\tau & ::= & \{\bar{\sigma}\} \mid \{\bar{r}\} \qquad\qquad\qquad\qquad\qquad\qquad \textit{types} \\
\\
\sigma & ::= & \texttt{def}\ m(x : \tau) : \tau\ \texttt{with}\ \varepsilon\ \textit{labeled decls.} \\
\\
\Gamma & ::= & \varnothing \\
& | & \Gamma,\ x : \tau
\end{array}
$$

Notes:

- All declarations (σ-terms) are annotated by what effects they have.
- All methods take exactly one argument. If a method specifies no argument the argument is assumed to be of type `Unit`.
- The type $\{\bar{r}\}$ is a set of resources; there will only be one actual resource at run-time, and it will be one of the resources in the set. This covers the case where e.g. a conditional returns a different resource on either branch.

2.2 Static Semantics

$\boxed{\Gamma \vdash e : \tau \texttt{ with } \varepsilon}$

$$\frac{}{\Gamma,\ x : \tau \vdash x : \tau \texttt{ with } \varnothing}\ (\varepsilon\text{-VAR}) \qquad \frac{}{\Gamma,\ r : \{\bar{r}\} \vdash r : \{\bar{r}\} \texttt{ with } \varnothing}\ (\varepsilon\text{-RESOURCE})$$

$$\frac{\Gamma,\ x : \{\bar{\sigma}\} \vdash \overline{\sigma = e}\ \texttt{OK}}{\Gamma \vdash \texttt{new}_\sigma\ x \Rightarrow \overline{\sigma = e} : \{\bar{\sigma}\} \texttt{ with } \varnothing}\ (\varepsilon\text{-NEWOBJ}) \qquad \frac{\Gamma \vdash e_1 : \{\bar{r}\} \texttt{ with } \varepsilon_1}{\Gamma \vdash e_1.\pi : \texttt{Unit with } \{\bar{r}.\pi\} \cup \varepsilon_1}\ (\varepsilon\text{-OPERCALL})$$

$$\frac{\Gamma \vdash e_1 : \{\bar{\sigma}\} \texttt{ with } \varepsilon_1 \quad \Gamma \vdash e_2 : \tau_2 \texttt{ with } \varepsilon_2 \quad \sigma_i = \texttt{def } m_i(y : \tau_2) : \tau_3 \texttt{ with } \varepsilon_3}{\Gamma \vdash e_1.m_i(e_2) : \tau_3 \texttt{ with } \varepsilon_1 \cup \varepsilon_2 \cup \varepsilon_3}\ (\varepsilon\text{-METHCALL}_\sigma)$$

$\boxed{\Gamma \vdash \sigma = e\ \texttt{OK}}$

$$\frac{\Gamma,\ y : \tau_2 \vdash e : \tau_3 \texttt{ with } \varepsilon_3 \quad \sigma = \texttt{def } m(y : \tau_2) : \tau_3 \texttt{ with } \varepsilon_3}{\Gamma \vdash \sigma = e\ \texttt{OK}}\ (\varepsilon\text{-VALIDIMPL}_\sigma)$$

Notes:

- In ε-VAR, ε-RESOURCE, and ε-NEWOBJ the consequent has an expression typed with no effect. In these rules we may be gaining an authority for an effect but we must use it in some code for that effect to happen.
- ε-VALIDIMPL says that the return type and effects of the body of a method must be exactly the same as its declaration.

2.3 Dynamic Semantics

$\boxed{e \longrightarrow e \mid \varepsilon}$

$$\frac{e_1 \longrightarrow e_1' \mid \varepsilon}{e_1.m(e_2) \longrightarrow e_1'.m(e_2) \mid \varepsilon}\ (\text{E-METHCALL1}_\sigma) \qquad \frac{v_1 = \texttt{new}_\sigma\ x \Rightarrow \overline{\sigma = e} \quad e_2 \longrightarrow e_2' \mid \varepsilon}{v_1.m(e_2) \longrightarrow v_1.m(e_2') \mid \varepsilon}\ (\text{E-METHCALL2}_\sigma)$$

$$\frac{v_1 = \texttt{new}_\sigma\ x \Rightarrow \overline{\sigma = e} \quad \texttt{def m}(y : \tau_1) : \tau_2 \texttt{ with } \varepsilon = e \in \overline{\sigma = e}}{v_1.m(v_2) \longrightarrow [v_1/x, v_2/y]e \mid \varnothing}\ (\text{E-METHCALL3}_\sigma)$$

$$\frac{e_1 \longrightarrow e_1' \mid \varepsilon}{e_1.\pi \longrightarrow e_1'.\pi \mid \varepsilon}\ (\text{E-OPERCALL1}) \qquad \frac{}{r.\pi \longrightarrow \texttt{unit} \mid \{r.\pi\}}\ (\text{E-OPERCALL2})$$

$\boxed{e \longrightarrow_* e \mid \varepsilon}$

$$\frac{}{e \longrightarrow_* e \mid \varnothing}\ (\text{E-MULTISTEP1}) \qquad \frac{e \longrightarrow e' \mid \varepsilon}{e \longrightarrow_* e' \mid \varepsilon}\ (\text{E-MULTISTEP2})$$

$$\frac{e \longrightarrow_* e' \mid \varepsilon_1 \quad e' \longrightarrow_* e'' \mid \varepsilon_2}{e \longrightarrow_* e'' \mid \varepsilon_1 \cup \varepsilon_2}\ (\text{E-MULTISTEP3})$$

Notes:

- A multi-step involves *zero* or more applications of a small-step.
- Multi-step rules accumulate the run-time effects produced by the individual small-steps.
- The only rule which produces effects is E-OperCall2 (the rule for evaluating operations on resources).
- Method calls are evaluated by performing substitution on the body of the method. There is no store.

2.4 Substitution Function

Definition 2.4.1. (Substitution)
$[e'/z]e$ means 'substitute every free occurrence of z in e for e'. Here's a definition over rules in the grammar.

- $[e'/z]z = e'$
- $[e'/z]y = y$, if $y \neq z$
- $[e'/z]r = r$
- $[e'/z](e_1.m(e_2)) = ([e'/z]e_1).m([e'/z]e_2)$
- $[e'/z](e_1.\pi) = ([e'/z]e_1).\pi$
- $[e'/z](\texttt{new}_\sigma\ x \Rightarrow \overline{\sigma = e}) = \texttt{new}_\sigma\ x \Rightarrow \overline{\sigma = [e'/z]e}$, if $z \neq x$ and $z \notin \texttt{freevars}(e_i)$

We use the convention of *alpha-conversion* to make substitution capture-avoiding. In practice, this means that whenever substitution is undefined because of variable capture, we rename variables to meet the side conditions (see Benjamin Pierce, "Types and Programming Languages" p. 71 for more details).

Substitution of multiple variables is written $[e_1/z_1, ..., e_n/z_n]e$, which is shorthand for $[e_n/z_n]...[e_1/z_1]e$

2.5 Soundness Theorem

In this section we build up to the soundness theorem for our typing system. We do this by proving progress and two preservation theorems about types and static effects. The two preservation theorems directly give us the soundness statement.

Lemma 2.5.1. (Canonical Forms)

Statement. Suppose e is a value. The following are true:

- If $\Gamma \vdash e : \{\bar{r}\}\ \texttt{with}\ \varepsilon$, then $e = r$ for some $r \in R$.
- If $\Gamma \vdash e : \{\bar{\sigma}\}\ \texttt{with}\ \varepsilon$, then $e = \texttt{new}_\sigma\ x \Rightarrow \overline{\sigma = e}$.

Furthermore, $\varepsilon = \varnothing$ in each case.

Proof. These typing judgements each appear exactly once in the conclusion of different rules. The result follows by inversion of ε-Resource and ε-NewObj respectively.

□

Lemma 2.5.2. (Substitution Lemma)

Statement. If $\Gamma, z : \tau'\ \texttt{with}\ \varepsilon' \vdash e : \tau\ \texttt{with}\ \varepsilon$, and $\Gamma \vdash e' : \tau'\ \texttt{with}\ \varepsilon'$, then $\Gamma \vdash [e'/z]e : \tau\ \texttt{with}\ \varepsilon$.

Intuition If you substitute z for something of the same type, the type of the whole expression stays the same after substitution.

Proof. By structural induction on possible derivations of $\Gamma \vdash e : \tau\ \texttt{with}\ \varepsilon$. First, if z does not appear in e then $[e'/z]e = e$, so the statement holds vacuously. So without loss of generality we assume z appears somewhere in e and consider the last rule used in the derivation, and then the location of z.

Case. ε-Var.
Then $[e'/z]z = e_1$. By assumption $\Gamma \vdash e' : \tau\ \texttt{with}\ \varepsilon$, so $\Gamma \vdash [e'/z]z = e$.

Case. ε-Resource.
Then $e = r$. $\texttt{freevars}(r) = \varnothing$, so the statement holds vacuously.

Case. ε-NewObj.
Then $e = \texttt{new}_\sigma\ x \Rightarrow \overline{\sigma = e}$. z appears in some method body e_i. By inversion we know $\Gamma, x : \{\bar{\sigma}\} \vdash \overline{\sigma = e}\ \texttt{OK}$. The only rule with this conclusion is ε-ValidImpl$_\sigma$; by inversion on that we have:
- $\sigma_i = \texttt{def}\ m_i(y : \tau_1) : \tau_2\ \texttt{with}\ \varepsilon$
- $\Gamma, y : \tau_1 \vdash e_i : \tau_2\ \texttt{with}\ \varepsilon$

$\Gamma, z : \tau \vdash \sigma_i = e_i\ \texttt{OK}$ via the inductive assumption. We can use ε-ValidImpl$_\sigma$ to conclude $\overline{\sigma = [e'/z]e}\ \texttt{OK}$. Then by ε-NewObj we type $[e'/z]e$ to the same as the original.

Case. ε-OperCall.
Then $e = e_1.\pi$. The variable z must appear in e_1. By rule inversion we have a sub-derivation for the type of both sub-expressions so applying the inductive assumption we know e_1 and $[e'/z]e_1$ have the same types. Since e and $[e'/z]e$ have the same syntactic structure, and their corresponding subexpressions have the same types, then ε-OperCall will type $[e'/z]e$ to the same thing as e.

Case. ε-MethCall$_\sigma$.
Then $e = e_1.m_i(e_2)$. The variable z must appear in e_1 or e_2. By rule inversion we have a sub-derivation for both so applying the inductive hypothesis we know the types of e_1 and $[e'/z]e_1$ are the same, and the types of e_2 and $[e'/z]e_2$ are the same. Since e and $[e'/z]e$ have the same syntactic structure, and their corresponding subexpressions have the same types, then ε-MethCall types $[e'/z](e_1.m_i(e_2))$ to the same as $e_1.m_i(e_2)$.
□

Corollary. If $\Gamma, z_i : \tau_i' \vdash e : \tau\ \texttt{with}\ \varepsilon$, and $\Gamma \vdash e_i' : \tau_i'\ \texttt{with}\ \varepsilon_i'$, then $\Gamma \vdash [e_1'/z_1, ..., e_n'/z_n]e : \tau\ \texttt{with}\ \varepsilon$. This follows by the definition $[e_1'/z_1, ..., e_n'/z_n]e = [e_n'/z_n]...[e_1'/z_1]e$ and induction on the length n.

Theorem 2.5.3. (Progress Theorem)

Statement. If $\Gamma \vdash e_A : \tau_A\ \texttt{with}\ \varepsilon_A$, either e_A is a value or a small-step $e_A \longrightarrow e_B \mid \varepsilon$ can be applied.

Proof. By induction on possible derivations of $\Gamma \vdash e_A : \tau_A\ \texttt{with}\ \varepsilon_A$. Consider the last rule used.

Case. ε-Var, ε-Resource, ε-NewObj.
Then e_A is a value.

Case. ε-MethCall$_\sigma$.
Then $e_A = e_1.m_i(e_2)$ and the following are known:
- $e_A : \tau_3\ \texttt{with}\ \varepsilon_1 \cup \varepsilon_2 \cup \varepsilon_3$
- $e_1 : \{\overline{\sigma}\}\ \texttt{with}\ \varepsilon_1$
- $e_2 : \tau_2\ \texttt{with}\ \varepsilon_2$
- $\sigma_i = \texttt{def}\ m_i(y : \tau_2) : \tau_3\ \texttt{with}\ \varepsilon_3$

We look at the cases for when e_1 and e_2 are values.

Subcase. e_1 is not a value. The derivation of $e_A : \tau\ \texttt{with}\ \varepsilon_A$ includes the subderivation $e_1 : \{\bar{\sigma}\}\ \texttt{with}\ \varepsilon_1$. By the inductive hypothesis $e_1 \longrightarrow e_1' \mid \varepsilon$. Then E-MethCall1 gives the reduction $e_A \longrightarrow e_1'.m_i(e_2) \mid \varepsilon$.

Subcase. e_2 is not a value. Without loss of generality, $e_1 = v_1$ is a value. Also, $e_2 : \tau_2\ \texttt{with}\ \varepsilon_2$ is a subderivation. By inductive hypothesis, $e_2 \longrightarrow e_2' \mid \varepsilon$. Then E MethCall2$_\sigma$ gives the reduction $e_A \longrightarrow v_1.m_i(e_2') \mid \varepsilon$.

Subcase. $e_1 = v_1$ and $e_2 = v_2$ are values. By Canonical Forms, $e_1 = \texttt{new}_\sigma\ x \Rightarrow \overline{\sigma = e}$. Also, $\texttt{def}\ m_i(y : \tau_2) : \tau_3\ \texttt{with}\ \varepsilon_3 = e_i \in \overline{\sigma = e}$. By the assumption of the typing rule used, the receiver and argument are well-typed for the method m_i. Then E-MethCall3$_\sigma$ gives the reduction $e_1.m_i(e_2) \longrightarrow [v_1/x, v_2/y]e_i \mid \varnothing$.

Case. ε-OperCall.
Then $e_A = e_1.\pi$ and the following are known:

- $e_A : \texttt{Unit with } \{r.\pi\} \cup \varepsilon_1$
- $e_1 : \{\bar{r}\} \texttt{ with } \varepsilon_1$

We look at the cases for when e_1 is a value.

Subcase. e_1 is not a value. Then $e_1 : \{\bar{r}\} \texttt{ with } \varepsilon_1$ is a subderivation. Applying inductive assumption, we have $e_1 \longrightarrow e_1' \mid \varepsilon$. Then E-OperCall1 gives the reduction $e_1.\pi \longrightarrow e_1'.\pi \mid \varepsilon$.

Subcase. e_1 is a value. By Canonical Forms, $\exists r \in R \mid e_1 = r$. Then E-OperCall3 gives the reduction $r.\pi \longrightarrow \texttt{unit} \mid \{r.\pi\}$.

□

Theorem 2.5.4. (Preservation Theorem)

Statement. Suppose the following hold:

- $\Gamma \vdash e_A : \tau_A \texttt{ with } \varepsilon_A$
- $e_A \longrightarrow e_B \mid \varepsilon$

Then $\Gamma \vdash e_B : \tau_B \texttt{ with } \varepsilon_B$, where $\tau_B = \tau_A$. Also, $\varepsilon \subseteq \varepsilon_A$ and $\varepsilon_B \subseteq \varepsilon_A$ and $\forall r.\pi \in \varepsilon_A \setminus \varepsilon_B \mid r.\pi \in \varepsilon$.

Intuition. Reduction preserves the relevant static effects in the sense that you only lose static effects during a computation if they actually happen. So you can't gain static effects during reduction, and every lost static effect is "accounted for".

Proof. By induction on possible derivations of $\Gamma \vdash e_A : \tau_A \texttt{ with } \varepsilon_A$. Consider the last rule used.

Case. ε-Resource, ε-Var, ε-NewObj.
e_A is a value so no reduction can be applied to it. The theorem statement is vacuously satisfied.

Case. ε-MethCall$_\sigma$.
Then $e_A = e_1.m_i(e_2)$ and the following are true:
- $\Gamma \vdash e_A : \tau_3 \texttt{ with } \varepsilon_1 \cup \varepsilon_2 \cup \varepsilon_3$
- $\Gamma \vdash e_1 : \{\overline{\sigma}\} \texttt{ with } \varepsilon_1$
- $\Gamma \vdash e_2 : \tau_2 \texttt{ with } \varepsilon_2$
- $\sigma_i = \texttt{def } m_i(y : \tau_2) : \tau_3 \texttt{ with } \varepsilon_3$

The type to be preserved is τ_3. We do a case analysis on the reductions applicable to $e_1.m_i(e_2)$.

Subcase. E-MethCall1$_\sigma$. Then $e_1 \longrightarrow e_1' \mid \varepsilon$ and $e_B = e_1'.m_i(e_2)$. By inductive assumption $\Gamma \vdash e_1' : \{\overline{\sigma}\} \texttt{ with } \varepsilon_1'$. The type of e_1' is the same as e_1, and the other subexpressions in e_B are identical to e_A. By applying ε-MethCall we have $\Gamma \vdash e_1'.m_i(e_2) : \tau_3 \texttt{ with } \varepsilon_1' \cup \varepsilon_2 \cup \varepsilon_3$. Then $\varepsilon_B = \varepsilon_1' \cup \varepsilon_2 \cup \varepsilon_3$ and $\varepsilon_A \setminus \varepsilon_B = \varepsilon_1 \setminus \varepsilon_1'$. For every $r.\pi \in \varepsilon_1 \setminus \varepsilon_1'$ we know $r.\pi \in \varepsilon$ by the inductive assumption. Since e_B has type τ_3 the type is preserved.

Subcase. E-MethCall2$_\sigma$. Then $e_1 = v_1 = \texttt{new}_\sigma\ x \Rightarrow \overline{\sigma = e}$, and $e_2 \longrightarrow e_2' \mid \varepsilon$ and $e_B = v_1.m_i(e_2')$. By inductive assumption $\Gamma \vdash e_2' : \tau_2 \texttt{ with } \varepsilon_2$. The type of e_2' is the same as e_2, and the other subexpressions in e_B are identical to e_A. By applying ε-MethCall we have $\Gamma \vdash e_B = v_1.m_i(e_2) : \tau_3 \texttt{ with } \varepsilon_1 \cup \varepsilon_2' \cup \varepsilon_3$. Then $\varepsilon_B = \varepsilon_1 \cup \varepsilon_2' \cup \varepsilon_3$ and $\varepsilon_A \setminus \varepsilon_B = \varepsilon_2 \setminus \varepsilon_2'$. For every $r.\pi \in \varepsilon_2 \setminus \varepsilon_2'$ we know $r.\pi \in \varepsilon$ by the inductive assumption. Since e_B has the type τ_3 the type is preserved.

Subcase. E-MethCall3$_\sigma$ Then $e_1 = v_1 = \texttt{new}_\sigma\ x \Rightarrow \overline{\sigma = e}$, and $\texttt{def } m_i(y : \tau_2) : \tau_3 \texttt{ with } \varepsilon_3 = e' \in \overline{\sigma = e}$, and $e_2 = v_2$ is a value, and $v_1.m_i(v_2) \longrightarrow [v_1/x, v_2/y]e' \mid \varnothing$.

We already know $e_1 : \{\overline{\sigma}\} \texttt{ with } \varepsilon_1$. The only rule with this conclusion is ε-NewObj. By inversion, $\overline{\sigma = e}$ OK. The only rule with this conclusion is ε-ValidImpl$_\sigma$. By inversion, $\Gamma, y : \tau_2 \vdash e' : \tau_3 \texttt{ with } \varepsilon_3$.

Now $e_B = [v_1/x, v_2/y]e'$ because the rule E-MethCall3 was used. Since $\Gamma, y : \tau_2 \vdash e' : \tau_3 \texttt{ with } \varepsilon_3$, and $\Gamma \vdash v_2 : \tau_2$ With ε_2, then by the substitution lemma, $\Gamma \vdash [v_2/y]e' : \tau_3 \texttt{ with } \varepsilon_3$.

We can strengthen the Γ in that last statement to obtain $\Gamma, x : \{\bar{\sigma}\} \texttt{ with } \varnothing \vdash [v_2/y]e' : \tau_3 \texttt{ with } \varepsilon_3$. By assumption we know $\Gamma \vdash v_1 : \{\overline{\sigma}\} \texttt{ with } \varnothing$. By another application of the substitution lemma,

$\Gamma \vdash [v_1/x, v_2/y]e' : \tau_3$ `with` ε_3.

(But this seems fishy: shouldn't the Γ in the premise of the ε-ValidImpl$_\sigma$ contain some mention of x? otherwise an object can't invoke methods on itself.

So $e_B : \tau_3$ `with` ε_3, which means $\varepsilon_B = \varepsilon_3$. Since $e_1 = v_1$ and $e_2 = v_2$ are values, by Canonical Forms $\varepsilon_1 = \varepsilon_2 = \varnothing$. So $\varepsilon_A = \varepsilon_3$. Then $\varepsilon_A \setminus \varepsilon_B = \varnothing$, so there are no lost effects to account for. Since e_B has the type τ_3 the type is preserved.

Case. ε-OperCall.
Then $e_A = e_1.\pi$: `Unit` , and we know:
- $\Gamma \vdash e_A$: `Unit with` $\{r, \pi\} \cup \varepsilon_1$
- $\Gamma \vdash e_1 : \{\bar{r}\}$ `with` ε_1

The type to be preserved is `Unit`. There are two reduction rules applicable to terms of the form $e_1.\pi$.

Subcase. E-OperCall1. Then $e_1 \longrightarrow e_1' \mid \varepsilon$ and $e_B = e_1'.\pi$. By inductive assumption, $\Gamma \vdash e_1' : \{\bar{r}\}$ `with` ε_1'. The type of e_1 is the same as e_1'. By applying ε-OperCall we have $\Gamma \vdash e_B = e_1'.\pi$: `Unit with` $\{r.\pi\} \cup \varepsilon_1'$. Then $\varepsilon_B = \{r.\pi\} \cup \varepsilon_1'$ and $\varepsilon_A \setminus \varepsilon_B = \varepsilon_1 \setminus \varepsilon_1'$. For every $r.\pi \in \varepsilon_1 \setminus \varepsilon_1'$ we know $r.\pi \in \varepsilon$ by the inductive assumption. Since e_B has the type `Unit` the type is preserved.

Subcase. E-OperCall2. Then $r.\pi \longrightarrow$ `unit` $\mid \{r.\pi\}$ and $e_B =$ `unit`. By Canonical Forms, $\varepsilon_1 = \varnothing$, so we can more specifically state the typing judgement as $\Gamma \vdash e_A$: `Unit with` $\{r.\pi\}$. By a degenerate case of ε-NewObj, $\Gamma \vdash$ `unit` : `Unit with` $\varnothing$. That means $\varepsilon_B = \varnothing$ so then $\varepsilon_A \setminus \varepsilon_B = \{r.\pi\}$, which is exactly the set ε of runtime effects; our only lost effect is accounted for. Since $e_B =$ `unit` has the type `Unit` the type is preserved.

□