Semantics for Exception Evaluation

서울대학교 전기·정보공학부 이재호, 이준협

December 29, 2022

1 Semantic domains

σ	$\in$	Env	$=$	$Id \to Val$	environment
M	$\in$	Mem	$=$	$Loc \to Val$	memory
v	$\in$	Val	$=$	$Prim + Closure + Arr + Lbl$	values
π	$\in$	$Prim$	$=$	$\{+, -, \texttt{raise}, \cdots\}$	primitive operators
c	$\in$	$Const$	$=$	$\mathbb{Z} + \mathbb{R} + \mathbb{B} + \cdots$	contants
x	$\in$	Id	$=$	identifiers in $\wp$	identifiers
f	$\in$	$Closure$	$=$	$Expr \times Env$	functions
$[\ell_1; \dots; \ell_m]$	$\in$	Arr	$=$	Loc^*	array-like data (arrays, records, modules, etc.)
κ	$\in$	$Ctor$	$=$	κ in $\wp$	constructors
$\langle \kappa, [\ell_1; \dots; \ell_m] \rangle$	$\in$	Lbl	$=$	$Ctor \times Arr$	labeled data (variants, exns, etc.)
$\underline{\langle \kappa, [\ell_1; \dots; \ell_m] \rangle}$	$\in$	$\underline{Lbl}$	$=$	Lbl	raised exceptions
ℓ	$\in$	Loc			location

expression	e	$\in$	$Expr$	
	e	$::=$	π	primitive operator
		$\mid$	c	constant
		$\mid$	x	id
		$\mid$	$\kappa\ e$	labeled data
		$\mid$	$\lambda\ (p\ \ e)^+$	function
		$\mid$	$e\ e$	application
		$\mid$	$[e^*]$	array-like data
pattern	p	$::=$	p_v	value pattern
		$\mid$	$\underline{p_v}$	"computation" pattern
value pattern	p_v	$::=$	$\_$	wildcard
		$\mid$	x	variable
		$\mid$	$\kappa\ p$	labeled pattern
		$\mid$	$[p^*]$	array-like data
		$\mid$	c	constant

$$\mathscr{E} : \underbrace{Expr \to Env \times Mem \to (Val + Exn) \times Mem}_{D}$$
$$\mathscr{F} : D \to D$$
$$\mathscr{E} = \text{fix}\,\mathscr{F}$$

1.1 One *Arr* to rule them all

1.1.1 Custom record

```
type myty = {a: int, mutable b: int}
let x = {a: 0, b: 0}
let y = x
y.b = 1
// x.b & y.b both 1
```

$$\sigma(\texttt{x}) = \ell_x$$
$$\sigma(\texttt{y}) = \ell_y$$
$$s(\ell_{\texttt{x}}) = [\ell_{\texttt{a}}, \ell_{\texttt{b}}]$$
$$s(\ell_{\texttt{y}}) = [\ell_{\texttt{a}}, \ell_{\texttt{b}}]$$
$$s(\ell_{\texttt{a}}) = 0$$
$$s(\ell_{\texttt{b}}) = 0 \rightsquigarrow 1$$

1.1.2 Exception reference

```
exception Exn1(int)
exception Exn2
let x = ref(Exn1 0)
let y = x
y := Exn2
```

$$\sigma(\texttt{x}) = \ell_x$$
$$\sigma(\texttt{y}) = \ell_y$$
$$s(\ell_{\texttt{x}}) = s(\ell_{\texttt{y}}) = [\ell_{\texttt{contents}}]$$
$$s(\ell_{\texttt{contents}}) = \langle \texttt{Exn1}, [\ell_\iota] \rangle \rightsquigarrow \langle \texttt{Exn2}, [] \rangle$$
$$s(\ell_\iota) = 0$$

1.1.3 Module

```
module type MySig = {
  let a: int
}
module MyMod = {
  let a = 0
}
module MyFun = (M: MySig) => {
  let b = M.a + 1
}
module M = MyFun(MyMod)
let {b: x} = module(M)
```

$$\sigma(\texttt{MyMod}) = \ell_{\texttt{MyMod}}$$
$$\sigma(\texttt{M}) = \ell_{\texttt{M}}$$
$$\sigma(\texttt{x}) = \ell_{\texttt{x}}$$
$$s(\ell_{\texttt{MyMod}}) = [\ell_{\texttt{a}}]$$
$$s(\ell_{\texttt{M}}) = [\ell_{\texttt{b}}]$$
$$s(\ell_{\texttt{x}}) = 1$$
$$s(\ell_{\texttt{a}}) = 0$$
$$s(\ell_{\texttt{b}}) = 1$$

```
var MyMod = {a: 0};
function MyFun(M) {
  var b = M.a + 1 | 0;
  return {b: b};
}
var b = 1;
var M = {b: b};
var x = b;
```

2 $\mathscr{C}$ to $\mathscr{G}$

We first define the set expressions that represent sets of values an arbitrary expression can have. Note that the set expressions are divided into two kinds, one for propagating function applications and another for propagating filtered patterns.

"Value" set expressions	v	$::=$	$\top$	*top*
		$\mid$	c	*constant*
		$\mid$	$\mathscr{X}$	*set variable*
		$\mid$	π	*primitive operator*
		$\mid$	$\lambda x.e$	*function*
		$\mid$	$\pi_{v\mid p}[(-\mid\mathscr{X})^*]$	*primitive application*
		$\mid$	$\mathsf{app}_{v\mid p}(\mathscr{X}, [(-\mid\mathscr{X})^*])$	*function application*
		$\mid$	$\langle(-\mid\kappa), [\ell^*]\rangle$	*variants, records, tuples, arrays*
		$\mid$	$\mathsf{fld}(\mathscr{X}, (-\mid\kappa), i)$	*field*
		$\mid$	$\mathscr{X} - p$	*pattern filtering*
"Pattern" set expressions	p	$::=$	$\hat{\top}$	*top*
		$\mid$	$\hat{c}$	*constant*
		$\mid$	$\langle(-\mid\kappa), [p^*]\rangle$	*variants, records, tuples, arrays*
		$\mid$	$\hat{\ell}$	*location*
		$\mid$	$\hat{\ell}\mid_p$	*location constrained by p*

Next the structure of possible set constraints are illustrated. The constraints collect what each set variable will contain, what each update will do, and what each location will contain.

Constraints $\mathscr{C}$	$\mathscr{X}$	$\supseteq$	v	*set variables*
	$\mathsf{fld}(\mathscr{X}, -, i)$	$\supseteq$	v	*updates*
	$!\ell$	$\supseteq$	v	*variant, record, tuple, array elements*
Grammar $\mathscr{G}$	$\hat{\mathscr{X}}$	$\supseteq$	p	*set variables*
	$!\hat{\ell}$	$\supseteq$	p	*variant, record, tuple, array elements*

Note that $\langle\_, [\ell^*]\rangle$ is used to represent the *Arr* datatype in section 1.

We need to find the least fixpoint $\mathsf{lfp}(\lambda(C,G).\mathscr{F}(\mathscr{C} \cup C, G)) =: (\mathscr{C}', \mathscr{G})$. $\mathscr{F}(C,G)$ takes a set of constraints C and a grammar G and performs "one step of resolution" to return a partially-resolved (C', G'). $\mathscr{C}$ is the initial set of constraints obtained from the program.

What we want is a good definition of $\mathscr{F}$ so that $(C,G) \sqsubseteq \mathscr{F}(C,G)$ and $\mathscr{F}^\infty(C,G)$ converges surely while safely approximating all possible values. $\bot_C$ and $\bot_G$ is defined as the constraint/grammar only having the production to $\bot$(representing an empty set). Note that a production in G specifies some pattern that $\mathscr{X}$ and $!\hat{\ell}$ might match.

2.1 Definition of $\mathcal{F}$

$\mathcal{F}(C, G)$: Look at the "productions" in C, determine what can be added to C and G.

Preliminaries:

$$\mathsf{len}(l) := \begin{cases} 0 & (l = []) \\ \mathsf{len}(\mathsf{tl}(l)) & (\mathsf{hd}(l) = \_) \\ \mathsf{len}(\mathsf{tl}(l)) + 1 & (\mathsf{hd}(l) \neq \_) \end{cases}$$

$$\mathsf{merge}(l, l') := \begin{cases} l' & (l = []) \\ l & (l' = []) \\ \mathsf{hd}(l') :: \mathsf{merge}(\mathsf{tl}(l), \mathsf{tl}(l')) & (\mathsf{hd}(l) = \_) \\ \mathsf{hd}(l) :: \mathsf{merge}(\mathsf{tl}(l), l') & (\mathsf{hd}(l) \neq \_) \end{cases}$$

That is, $\mathsf{merge}(l, l')$ is performed by plugging in elements of l' one by one into the _ places of l(including $\_ \in l'$), then concatenating the rest of l' to the tail side of l if there is no more free space.

Now, let's define $\mathcal{F}$:

1. For productions $\mathcal{X}|!\ell \supseteq \top \mid c \mid | \langle(\_|\kappa), [\ell^*]\rangle$, add the same productions "with a hat" to G if they are not already in G.

2. For production $\mathcal{X}|!\ell \supseteq \mathcal{X}_1$,

    (a) If $\hat{\mathcal{X}}_1 \supseteq \star$ is in G, add $\hat{\mathcal{X}}|!\hat{\ell} \supseteq \star$ to G.

    (b) If $\mathcal{X}_1 \supseteq \pi \mid \lambda x.e \mid \mathsf{app}_v(\mathcal{X}_2, \_ :: \mathtt{tl}) \mid \pi_v\ \mathtt{l}$ is in C, add those to $\mathcal{X}|!\ell$ in C.

3. For production $\mathcal{X}|!\ell \supseteq \pi_v\ \mathtt{l}$, when $\mathsf{len}(\mathtt{l}) = \mathsf{arity}(\pi)$ and π is not `raise`, add $\hat{\mathcal{X}}|!\hat{\ell} \supseteq \top$ to G (constant propagation may be added). Note that `ignore`, `identity`, `reverse` must also be put into consideration, but this is trivial.

4. For production $\mathcal{X}|!\ell \supseteq \pi_p\ \mathtt{l}$, when $\mathsf{len}(\mathtt{l}) = \mathsf{arity}(\pi)$ and π is `raise`, add $\hat{\mathcal{X}}|!\hat{\ell} \supseteq \mathsf{hd}(\mathtt{l})$ to G.

5. For production $\mathcal{X}|!\ell \supseteq \mathsf{app}_v(\mathcal{X}_1, [])$, this only happens on a `Lazy.force`, so if $\mathcal{X}_1 \supseteq \lambda.e$ is in C, then add $\hat{\mathcal{X}}|!\hat{\ell} \supseteq \mathcal{X}(e)$ to G.

6. For production $\mathcal{X}|!\ell \supseteq \mathsf{app}_p(\mathcal{X}_1, [])$, this only happens on a `Lazy.force`, so if $\mathcal{X}_1 \supseteq \lambda.e$ is in C, then add $\hat{\mathcal{X}} \supseteq \mathcal{X}|!\hat{\ell}(\underline{e})$ to G.

7. For production $\mathcal{X}|!\ell \supseteq \mathsf{app}_v(\mathcal{X}_1, \mathcal{X}_2 :: \mathtt{tl})$,

    (a) If $\mathcal{X}_1 \supseteq \pi$ is in C, add $\mathcal{X}|!\ell \supseteq \pi_v\ \mathcal{X}_2 :: \mathtt{tl}$ to C.

    (b) If $\mathcal{X}_1 \supseteq \pi_v\ \mathtt{l}$ is in C, add $\mathcal{X}|!\ell \supseteq \pi_v\ \mathsf{merge}(\mathtt{l}, \mathcal{X}_2 :: \mathtt{tl})$ to C.

    (c) If $\mathcal{X}_1 \supseteq \lambda x.e$ is in C, $\mathtt{tl} \neq []$, add $\mathcal{X}|!\ell \supseteq \mathsf{app}_v(\mathcal{X}(e), \mathtt{tl})$, $\mathcal{X}(E_x) \supseteq \mathcal{X}_1$ to C.

    (d) If $\mathcal{X}_1 \supseteq \lambda x.e$ is in C, $\mathtt{tl} = []$, add $\mathcal{X}|!\ell \supseteq \mathcal{X}(e)$, $\mathcal{X}(E_x) \supseteq \mathcal{X}_1$ to C.

    (e) If $\mathcal{X}_1 \supseteq \mathsf{app}_v(\mathcal{X}_3, \_ :: \mathtt{tl}')$ is in C, add $\mathcal{X}|!\ell \supseteq \mathsf{app}_v(\mathcal{X}_3, \mathcal{X}_2 :: \mathsf{merge}(\mathtt{tl}, \mathtt{tl}'))$ to C.

8. For production $\mathcal{X}|!\ell \supseteq \mathsf{app}_p(\mathcal{X}_1, \mathcal{X}_2 :: \mathtt{tl})$,

    (a) If $\mathcal{X}_1 \supseteq \pi$ is in C, add $\mathcal{X}|!\ell \supseteq \pi_p\ \mathcal{X}_2 :: \mathtt{tl}$ to C.

(b) If $\mathcal{X}_1 \supseteq \pi_v$ $\mathtt{l}$ is in C, add $\mathcal{X}|!\ell \supseteq \pi_p$ $\mathsf{merge}(\mathtt{l}, \mathcal{X}_2 :: \mathtt{tl})$ to C.

(c) If $\mathcal{X}_1 \supseteq \lambda x.e$ is in C, add $\mathcal{X}|!\ell \supseteq \mathcal{X}(\underline{e})$, $\mathcal{X}(E_x) \supseteq \mathcal{X}_1$ to C. Additionally, if $\mathtt{tl} \neq []$ then also add $\mathsf{app}_p(\mathcal{X}(e), \mathtt{tl})$.

(d) If $\mathcal{X}_1 \supseteq \mathsf{app}_v(\mathcal{X}_3, \_ :: \mathtt{tl}')$ is in C, add $\mathcal{X}|!\ell \supseteq \mathsf{app}_p(\mathcal{X}_3, \mathcal{X}_2 :: \mathsf{merge}(\mathtt{tl}, \mathtt{tl}'))$ to C.

9. For production $\mathcal{X}|!\ell \supseteq \mathsf{fld}(\mathcal{X}_1, \_, i)$,

   (a) If $\hat{\mathcal{X}}_1 \supseteq \top$ is in G, add $\hat{\mathcal{X}}|!\hat{\ell} \supseteq \top$ to G.

   (b) If $\hat{\mathcal{X}}_1 \supseteq \langle(\_|\kappa), ...\hat{\ell}_i...\rangle$ is in G and if $!\hat{\ell}_i \supseteq$ something is in G, add $\hat{\mathcal{X}}|!\hat{\ell} \supseteq$ something to G.

   (c) If $\hat{\mathcal{X}}_1 \supseteq \langle(\_|\kappa), ...p_i...\rangle$ is in G, add $\hat{\mathcal{X}}|!\hat{\ell} \supseteq p_i$ to G.

10. For production $\mathcal{X}|!\ell \supseteq \mathsf{fld}(\mathcal{X}_1, \kappa, i)$,

    (a) If $\hat{\mathcal{X}}_1 \supseteq \top$ is in G, add $\hat{\mathcal{X}}|!\hat{\ell} \supseteq \top$ to G.

    (b) If $\hat{\mathcal{X}}_1 \supseteq \langle\kappa, ...\hat{\ell}_i...\rangle$ is in G and if $!\hat{\ell}_i \supseteq$ something is in G, add $\hat{\mathcal{X}}|!\hat{\ell} \supseteq$ something to G. The constructor κ must be matched.

    (c) If $\hat{\mathcal{X}}_1 \supseteq \langle\kappa, ...p_i...\rangle$ is in G, add $\hat{\mathcal{X}}|!\hat{\ell} \supseteq p_i$ to G.

11. For production $\mathsf{fld}(\mathcal{X}, \_, i) \supseteq$ something,

    (a) If $\hat{\mathcal{X}}_1 \supseteq \langle(\_|\kappa), ...\hat{\ell}_i...\rangle$ is in G, add $!\ell_i \supseteq$ something to C.

12. For production $\mathcal{X}|!\ell \supseteq \mathcal{X}_1 - p$, first define

$$
\mathsf{filter}(x, p) := \begin{cases}
\emptyset & (p = \top) \\
\emptyset & (x = p) \\
\{x\} & (x \neq p, p = c) \\
\{x\} & (x = \langle\kappa, \_\rangle, p = \langle\kappa', \_\rangle) \\
\mathsf{filter}(\langle\_, \underbrace{[\top; ...; \top]}_{n \text{ times}}\rangle, p) & (x = \top, p = \langle\_, [p_1; ...; p_n]\rangle) \\
\bigcup_{\hat{x} \supseteq y \in G} \mathsf{filter}(y, p) & ((x, \hat{x}) = (\hat{\mathcal{X}}, \hat{\mathcal{X}}) \text{ or } (\hat{\ell}, !\hat{\ell})) \\
\bigcup_{i=0}^{n-1} \left\langle \kappa, \left(\prod_{j=1}^{i} p_j\right) \mathsf{filter}(x_{i+1}, p_{i+1}) \left(\prod_{j=i+2}^{n} x_j\right) \right\rangle & \underbrace{(x = \langle\kappa, [x_1; ...; x_n]\rangle, p = \langle\kappa, [p_1; ...; p_n]\rangle)}_{\kappa \text{ may be } \_}
\end{cases}
$$

when $x \in \{p, \hat{\ell}, \hat{\mathcal{X}}\}$

Note that $\top$ happens on a π_v, and the type of p must match with the type of x, so we can reconstruct the type of $\top$ from p (when we use externals that return records). Also, in the list concatenation part of the last case, if the result of filter is $\emptyset$, the whole thing is $\emptyset$.

Now, add $\hat{\mathcal{X}}|!\hat{\ell} \supseteq \alpha$ for all $\alpha \in \mathsf{filter}(\hat{\mathcal{X}}_1, p)$ to G.

2.2 `Array.make` (resolution of π: maybe later)

For primitives π, a new constructor is generated, e.g., $\mathsf{app}_v(\pi, [])$ to $\mathsf{ctor}(\_, [\ell_{\text{new}}])$.

2.3 Brainstorming

$$\begin{aligned}
\langle\kappa, \ell_1\ell_2\ldots\ell_n\rangle - \langle\kappa, p_1p_2\ldots p_n\rangle &= \langle\kappa, (\ell_1 - p_1)\ell_2\ldots\ell_n\rangle \\
&\quad + \langle\kappa, p_1(\ell_2\ldots\ell_n - p_2\ldots p_n)\rangle \\
&= \ldots \\
&= \sum_{i=0}^{n-1}\left(\prod_{j=1}^{i} p_j\right)(\ell_{i+1} - p_{i+1})\left(\prod_{j=i+2}^{n} \ell_j\right) \\
\ell - p &=! \hat{\ell} - p
\end{aligned}$$

Here, the product notation stands for list concatenation, when p_i and ℓ_i stands for a single-element list. $\ell - p$ stands for the contents of the abstract location ℓ that is not matched by the pattern p : how to express this neatly?