

$\ell^*$	$::= \ell \mid \text{END}$	labels
$T$	$::= [t, \dots]$	stack
$t$	$::= \ell^* \times S$	stack frame
$S$	$::= [s, \dots]$	programs
$d$	$::= x = x \mid x = v \mid x = x(x)$	directives
$B$	$::= x \mapsto m, \dots$	bindings
$H$	$::= m \mapsto v, \dots$	heap
$v$	$::= \text{int}^+ \mid \text{int}^- \mid \text{int}^0 \mid \langle \eta, \text{fun}(x) \mapsto S \rangle$	values
$\eta$		pointers to scopes
$E$	$::= \eta \mapsto B$	environments
$P$	$::= \eta \mapsto \eta$	parental maps

LITERAL ASSIGNMENT

$$\begin{array}{c}
S(\ell) = \ell : \ell' : x = v \quad H' = H[m \mapsto v] \\
\hline
m \notin H \quad B = E(\eta) \quad B' = B[x \mapsto m] \quad E' = E[\eta \mapsto B'] \quad \ell \stackrel{s}{\blacktriangleleft} \ell'' \\
\hline
[\langle \ell, S \rangle] \parallel T, H, E, P, \eta \longrightarrow^1 [\langle \ell'', S \rangle] \parallel T, H', E', P, \eta
\end{array}$$

VARIABLE ASSIGNMENT

$$\begin{array}{c}
S(\ell) = \ell : \ell' : x_1 = v_2 \quad m = \text{LOOKUP}(\eta, P, E, x_1) \\
\hline
B = E(\eta) \quad B' = B[x_1 \mapsto m] \quad E' = E[\eta \mapsto B'] \quad \ell \stackrel{s}{\blacktriangleleft} \ell'' \\
\hline
[\langle \ell, S \rangle] \parallel T, H, E, P, \eta \longrightarrow^1 [\langle \ell'', S \rangle] \parallel T, H, E', P, \eta
\end{array}$$

FUNCTION CALL

$$\begin{array}{c}
S(\ell) = \ell : \ell' : x_1 = x_2(x_3) \quad m = (\eta, P, E, x_2) \\
H[m] = \langle \eta', \text{fun}(x_4) \mapsto S' \rangle \quad \eta'' \notin E \quad m' = (\eta, P, E, x_3) \\
B = \{x_4 \mapsto m'\} \quad E' = E[\eta'' \mapsto B] \quad P' = P \cup \{\eta'' \mapsto \eta'\} \quad S' = [\ell'' : \ell''' : d] \\
\hline
[\langle \ell, S \rangle] \parallel T, H, E, P, \eta \longrightarrow^1 [\langle \ell'', S' \rangle, \langle l, S \rangle] \parallel T, H', E', P', \eta''
\end{array}$$

Figure 1: Operation Semantics

## 1 CoPylot

### 1.1 Grammar

The grammar of the language to be analyzed appears in Figure 2.

We assume throughout the rest of this document that a fixed program  $\hat{S}$  is under analysis. *(TODO: describe here the idea of a bijection between labels and statements in this fixed program. – ZP)*

### 1.2 Control Flow

The grammar of control flow graphs appears in Figure 3. *(Discuss construction of initial graph. – ZP)*

$\hat{S} ::= [\hat{s}, \dots]$	<i>abstract programs</i>
$\hat{s} ::= \hat{\ell} : \hat{\ell} : \hat{d}$	<i>abstract statements</i>
$\hat{d} ::= \hat{x} = \hat{v} \mid \hat{x} = \hat{x}$	<i>abstract directives</i>
$\hat{v} ::= \mathbf{int}^+ \mid \mathbf{int}^- \mid \mathbf{int}^0$	<i>abstract values</i>
$\hat{x}$	<i>abstract variables</i>
$\hat{\ell}$	<i>abstract labels</i>

Figure 2: Normalized Python Language Grammar

$\hat{G} ::= \{\hat{g}, \dots\}$	<i>control flow graphs</i>
$\hat{g} ::= \hat{o} \blacktriangleleft^? \hat{o} \mid \hat{o} \ll \hat{o}$	<i>control flow graph edge</i>
$\hat{o} ::= \mathbf{START} \mid \mathbf{END} \mid \hat{s}$	<i>control flow graph nodes</i>

Figure 3: Control Flow Graph Grammar

We write  $\hat{o} \blacktriangleleft^? \hat{o}'$  to denote  $(\hat{o} \blacktriangleleft \hat{o}' \in \hat{G})$  when  $\hat{G}$  is understood from context. Likewise, we write  $\hat{o} \ll^? \hat{o}'$  to denote  $(\hat{o} \ll \hat{o}' \in \hat{G})$  when  $\hat{G}$  is understood from context.

We define a relation  $\longrightarrow^1$  to perform control flow graph closure.

**Definition 1.1.** *Let  $\hat{G} \longrightarrow^1 \hat{G}'$  be the least relation satisfying the rules appearing in Figure 4. Throughout these rules, the predicates  $\ll^?$  and  $\blacktriangleleft^?$  refer to graph  $\hat{G}$ .*

### 1.3 Value Lookup

The value lookup function uses the additional grammar in Figure 5.

**Definition 1.2.** *Given a control-flow graph  $\hat{G}$ , let  $\hat{G}(\hat{o}_0, \hat{K})$  be the function returning the least set  $\hat{V}$  which satisfies the following conditions:*

#### 1. Value Manipulation

<p>LEXICAL START</p> $\frac{\mathbf{START} \blacktriangleleft^? \hat{o}}{\hat{G} \longrightarrow^1 \hat{G} \cup \{\mathbf{START} \ll \hat{o}\}}$	<p>LITERAL ASSIGNMENT</p> $\frac{\hat{o}_1 = (\hat{x} = \hat{v}) \quad \hat{o}_1 \blacktriangleleft^? \hat{o}_2}{\hat{G} \longrightarrow^1 \hat{G} \cup \{\hat{o}_1 \ll \hat{o}_2\}}$
<p>VARIABLE ACCESSIBLE</p> $\frac{\hat{o}_1 = (\hat{x} = \hat{x}') \quad \hat{o}_1 \blacktriangleleft^? \hat{o}_2 \quad \hat{v} \in \hat{G}(\hat{o}_1, [\hat{x}']) \quad \hat{v} \neq \mathbf{UNDEFINED}}{\hat{G} \longrightarrow^1 \hat{G} \cup \{\hat{o}_1 \ll \hat{o}_2\}}$	

Figure 4: Control Flow Graph Closure

$$\begin{array}{ll}
\hat{K} & ::= [\hat{k}, \dots] & \text{lookup stacks} \\
\hat{k} & ::= \hat{x} \mid \hat{v} \mid \text{CAPTURE}(\mathbb{N}) \mid \text{JUMP}(\hat{o}) & \text{lookup stack elements}
\end{array}$$

Figure 5: Value Lookup Grammar

$$\begin{array}{l}
(a) \quad \boxed{\text{RESULT}} \\
\text{If } \hat{K} = [\hat{v}], \text{ then } \hat{v} \in \hat{V}.
\end{array}$$

## 2. Variable Lookup

$$\begin{array}{l}
(a) \quad \boxed{\text{VALUE DISCOVERY}} \\
\text{If } \hat{o}_1 \stackrel{?}{\ll} \hat{o}_0, \hat{o}_1 = \hat{\ell}_1 : \hat{\ell}_2 : \hat{x} = \hat{v}, \text{ and } \hat{K} = [\hat{x}] \parallel \hat{K}', \text{ then } \hat{G}(\hat{o}_1, [\hat{v}] \parallel \hat{K}') \subseteq \hat{V}. \\
(b) \quad \boxed{\text{VALUE SKIP}} \\
\text{If } \hat{o}_1 \stackrel{?}{\ll} \hat{o}_0, \hat{o}_1 = \hat{\ell}_1 : \hat{\ell}_2 : \hat{x}' = \hat{v}, \hat{K} = [\hat{x}] \parallel \hat{K}', \text{ and } \hat{x} \neq \hat{x}', \text{ then } \hat{G}(\hat{o}_1, \hat{K}) \subseteq \hat{V}. \\
(c) \quad \boxed{\text{VALUE ALIASING}} \\
\text{If } \hat{o}_1 \stackrel{?}{\ll} \hat{o}_0, \hat{o}_1 = \hat{\ell}_1 : \hat{\ell}_2 : \hat{x} = \hat{x}', \text{ and } \hat{K} = [\hat{x}] \parallel \hat{K}', \text{ then } \hat{G}(\hat{o}_1, [\hat{x}'] \parallel \hat{K}) \subseteq \hat{V}.
\end{array}$$