1.9 Language Semantics

The following sections contain the operational semantics of ROOPL++, as specified by syntax-directed inference rules.

1.9.1 Preliminaries

We define a memory location l to be a location. We define a location for integer variables to bind to a single location in program memory and a vector of memory locations for object and array variables, where the vector is the size of the object or array. A memory location is in the set of non-negative integers, \mathbb{N}_0 . An environment γ is a partial function mapping variables to memory locations. A store μ is a partial function mapping memory locations to values. An object is a tuple of a class name and an environment mapping fields to memory locations. A value is either an integer, an object or a memory location.

Applications of environments γ and stores μ are analogous to the type environment Γ , defined in section 1.8.1.

$$\begin{array}{ll} l \in \operatorname{Locs} &= \mathbb{N}_0 \\ \gamma \in \operatorname{Envs} &= \operatorname{VarIDs} \rightharpoonup \operatorname{Locs} \\ \mu \in \operatorname{Stores} &= \operatorname{Locs} \rightharpoonup \operatorname{Values} \\ & \operatorname{Objects} = \Big\{ \langle c_f, \ \gamma_f \rangle \mid c_f \in \operatorname{ClassIDs} \ \land \ \gamma_f \in \operatorname{Envs} \Big\} \\ v \in \operatorname{Values} &= \mathbb{Z} \cup \operatorname{Objects} \ \cup \operatorname{Locs} \end{array}$$

Figure 1.21: Semantic values, originally from [1]

1.9.2 Expressions

The judgment:

$$\langle \gamma, \mu \rangle \vdash_{expr} e \Rightarrow v$$

defines the meaning of expressions. We say that under environment γ and store μ , expression e evaluates to value v.

$$\frac{\langle \gamma, \mu \rangle \vdash_{expr} n \Rightarrow \bar{n}}{\langle \gamma, \mu \rangle \vdash_{expr} e_1 \Rightarrow v_1} \underbrace{\begin{array}{c} \langle \gamma, \mu \rangle \vdash_{expr} x \Rightarrow \mu \Big(\gamma(x) \Big) \end{array}}_{\langle \gamma, \mu \rangle} VAR \qquad \overline{\begin{array}{c} \langle \gamma, \mu \rangle \vdash_{expr} \mathbf{nil} \Rightarrow 0 \end{array}}_{\langle \gamma, \mu \rangle} NIL$$

$$\frac{\langle \gamma, \mu \rangle \vdash_{expr} e_1 \Rightarrow v_1 \qquad \langle \gamma, \mu \rangle \vdash_{expr} e_2 \Rightarrow v_2 \qquad [\![\otimes]\!] (v_1, v_2) = v}{\langle \gamma, \mu \rangle \vdash_{expr} e_1 \otimes e_2 \Rightarrow v} BINOP$$

Figure 1.22: Semantic inference rules for expressions, originally from [1]

As shown in figure 1.22, expression evaluation has no effects on the store. Logical values are represented by *truthy* and *falsy* values of any non-zero value and zero respectively. The evaluation of binary operators is presented in figure 1.24.

$$\frac{\langle \gamma, \mu \rangle \vdash_{expr} e \Rightarrow v \qquad \gamma(x) = l \qquad \mu(l)[v] = l' \qquad \mu(l') = w}{\langle \gamma, \mu \rangle \vdash_{expr} x[e] \Rightarrow w} \text{ArrElem}$$

Figure 1.23: Extension to the semantic inference rules for expression in ROOPL++

For ROOPL++, we extend the expression ruleset with a single rule for array element variables shown in figure 1.23. As with the expressions inference rules in ROOPL, this extension has no effect on the store.

Figure 1.24: Definition of binary expression operator evaluation, originally from [1]

1.9.3 Statements

The judgment

$$\gamma \vdash^{\Gamma}_{stmt} s : \mu \rightleftharpoons \mu'$$

defines the meaning of statements. We say that under environment γ , statement s with class map Γ reversibly transforms store μ to store μ' . Figure 1.25a, 1.25b and ?? defines the operational semantics of ROOPL++.

The inference rule SKIP defines the operational semantics of **skip** statements and has no effects on the store μ .

Rule SEQ defines statement sequences where the store potentially is updated between each statement execution.

$$\frac{\gamma \vdash_{stmt}^{\Gamma} \mathbf{skip} : \mu = \mu}{\gamma \vdash_{stmt}^{\Gamma} \mathbf{skip} : \mu = \mu} \overset{\text{SKIP}}{} \frac{\gamma \vdash_{stmt}^{\Gamma} \mathbf{s}_1 : \mu = \mu' \qquad \gamma \vdash_{stmt}^{\Gamma} \mathbf{s}_2 : \mu' = \mu''}{\gamma \vdash_{stmt}^{\Gamma} \mathbf{s}_1 \cdot \mathbf{s}_2 : \mu = \mu''}} \overset{\text{SEQ}}{} \frac{\gamma \vdash_{stmt}^{\Gamma} \mathbf{s}_1 \cdot \mathbf{s}_1 \cdot \mathbf{s}_2 : \mu = \mu''}{\gamma \vdash_{stmt}^{\Gamma} \mathbf{s}_1 \cdot \mathbf{s}_2 : \mu = \mu''} \overset{\text{SEQ}}{} \frac{\langle \gamma, \mu \rangle \vdash_{stmt}^{\Gamma} \mathbf{s}_1 \cdot \mathbf{s}_1 \cdot \mathbf{s}_2 \cdot \mu = \mu[\gamma(x)), \psi \lor_{st}']}{\gamma \vdash_{stmt}^{\Gamma} \mathbf{s}_1 \cdot \mathbf{s}_1 \cdot \mathbf{s}_2 \cdot \mu = \mu[\gamma(x_1) \mapsto \psi_2, \gamma(x_2) \mapsto \psi_1]} \overset{\text{SWPVAR}}{} \frac{\langle \gamma, \mu \rangle \vdash_{expr}^{\Gamma} \mathbf{s}_1 \cdot \mu = \mu' \qquad \gamma \vdash_{loop}^{\Gamma} (\mathbf{s}_1, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_2) : \mu' = \mu''}{\gamma \vdash_{stmt}^{\Gamma} \mathbf{from} \mathbf{s}_1 \cdot \mathbf{s}_1 \cdot \mu = \mu' \qquad \gamma \vdash_{loop}^{\Gamma} (\mathbf{s}_1, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_2) : \mu' = \mu''} \overset{\text{LoopMain}}{} \frac{\langle \gamma, \mu \rangle \vdash_{expr}^{\Gamma} \mathbf{s}_2 \Rightarrow 0 \qquad \gamma \vdash_{stmt}^{\Gamma} \mathbf{s}_1 \cdot \mu' = \mu'' \qquad \gamma \vdash_{loop}^{\Gamma} (\mathbf{s}_1, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_2) : \mu' = \mu''}{\gamma \vdash_{loop}^{\Gamma} (\mathbf{s}_1, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_2) : \mu = \mu''} \overset{\text{LoopRec}}{} \frac{\langle \gamma, \mu \rangle \vdash_{expr}^{\Gamma} \mathbf{s}_1 \Rightarrow 0 \qquad \gamma \vdash_{stmt}^{\Gamma} \mathbf{s}_1 \cdot \mu = \mu' \qquad \gamma \vdash_{loop}^{\Gamma} (\mathbf{s}_1, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_2) : \mu' = \mu''}{\gamma \vdash_{stmt}^{\Gamma} \mathbf{t}_1 \mathbf{t}_$$

Figure 1.25a: Semantic inference rules for statements, modified from [1]

Rule AssVar defines reversible assignment in which variable identifier x under environment γ is mapped to the value v' resulting in an updated store μ' . For variable swapping SWPVAR defines how value mappings between two variables are exchanged in the updated store.

For loops and conditionals, Rules LOOPMAIN, LOOPBASE and LOOPREC define the meaning of loop statements and IfTrue and IfFalse, similarly to the operational semantics of Janus, as presented in [3]. LOOPMAIN is entered if e_1 is true and each iteration enters LOOPREC until e_2 is false, in which case LOOPBASE is executed. Similarly, if e_1 and e_2 are true, rule IFTRUE is entered, executing the then-branch of the conditional. If e_1 and e_2 are false, the IFFALSE rule is

$$\gamma(this) = l \quad \gamma(x_0) = l \quad \mu(l) = l' \quad \mu(l') = \left\langle c, (l_1, \dots, l_m) \right\rangle \quad \gamma(y_i) = l_i'$$

$$fields \qquad methods$$

$$\Gamma(c) = \left\langle \overbrace{(x_1, \dots, x_m)}, \overbrace{(\dots, \mathbf{method} \ q(t_1 z_1, \dots, t_l z_k) \ s, \dots)} \right\rangle$$

$$\underline{\gamma' = [this \mapsto l, x_1 \mapsto l_1, \dots, l_m \mapsto v_m, z_1 \mapsto l_1', \dots, z_k \mapsto l_k'] \quad \gamma' \vdash_{stmt}^{\Gamma} \ s : \mu = \mu'}{\gamma \vdash_{stmt}^{\Gamma} \ \text{call} \ q(y_1, \dots, y_n) \ : \ \mu = \mu'} \quad \text{Call}$$

$$\frac{\gamma \vdash_{stmt}^{\Gamma} \ \text{call} \ q(y_1, \dots, y_n) \ : \ \mu' \Rightarrow \mu}{\gamma \vdash_{stmt}^{\Gamma} \ \text{uncall} \ q(y_1, \dots, y_n) \ : \ \mu = \mu'} \quad \text{UNCALL}$$

$$\gamma(x_0) = l \quad \mu(l) = l' \quad \mu(l') = \left\langle c, (l_1, \dots, l_m) \right\rangle \quad \gamma(y_i) = l_i'$$

$$\frac{fields}{\gamma \vdash_{stmt}^{\Gamma} \ \text{uncall} \ q(y_1, \dots, y_n) \ : \ \mu \Rightarrow \mu'} \quad \gamma' \vdash_{stmt}^{\Gamma} \ s : \mu \Rightarrow \mu'}$$

$$\gamma' = [this \mapsto l, x_1 \mapsto l_1, \dots, l_m \mapsto v_m, z_1 \mapsto l_1', \dots, z_k \mapsto l_k'] \quad \gamma' \vdash_{stmt}^{\Gamma} \ s : \mu \Rightarrow \mu'}{\gamma \vdash_{stmt}^{\Gamma} \ \text{call} \ x_0 :: q(y_1, \dots, y_n) \ : \ \mu \Rightarrow \mu'} \quad \text{CallObJ}$$

$$\frac{\gamma' \vdash_{stmt}^{\Gamma} \ \text{call} \ x_0 :: q(y_1, \dots, y_n) \ : \ \mu \Rightarrow \mu'}{\gamma \vdash_{stmt}^{\Gamma} \ \text{uncall} \ x_0 :: q(y_1, \dots, y_n) \ : \ \mu \Rightarrow \mu'} \quad \text{OBJUNCALL}}$$

$$\Gamma(c) = \left\langle \overbrace{(x_1, \dots, x_m)}, methods \right\rangle \quad \gamma' = \gamma[x \mapsto 0] \quad l_0 \notin \text{dom}(\mu) \ \dots \ l_m \notin \text{dom}(\mu)$$

$$\mu' = \mu \left[\gamma'(x) \mapsto l_0, l_0 \mapsto \left\langle c, (l_1, \dots, l_m) \right\rangle, l_1 \mapsto 0, \dots, l_m \mapsto 0 \right]$$

$$\frac{\gamma' \vdash_{stmt}^{\Gamma} \ \text{stm}}{\gamma \vdash_{stmt}^{\Gamma} \ \text{st}} \ : \ \mu' \Rightarrow \mu'' \quad \mu'' = \mu' \setminus \{\gamma'(x), l_0, \dots, l_m\}}{\gamma \vdash_{stm}^{\Gamma} \ \text{construct} \ c \ x \ s \ \text{destruct} \ x : \mu \Rightarrow \mu''} \quad \text{OBJBLOCK}$$

Figure 1.25b: Semantic inference rules for statements, modified from [1] (cont)

executed and the else-branch is executed.

Rules Call, Uncall, Callobj and Uncallobj respectively define local and non-local method invocations. For local methods, method q in current class c should be of arity n matching the number of arguments. The updated store μ' is obtained after statement body execution in the object environment. As local uncalling is the inverse of local calling, the direction of execution is simply reversed, and as such the input store a **call** statement serves as the output store of the **uncall** statement, similarly to techniques presented in [4, 3].

The statically scoped object blocks are defined in rule ObjBlock. The operation semantics of these blocks are similar to **local**-blocks from Janus. We add the reference x to a new environment and afterwards map location of x to the object tuple at location l_0 , containing the locations of

$$\begin{split} \langle \gamma, \mu \rangle \vdash^{\Gamma}_{stmt} & e_1 \Rightarrow v_1 \quad \langle \gamma, \mu \rangle \vdash^{\Gamma}_{stmt} & e_2 \Rightarrow v_2 \\ \\ \frac{\gamma(x) = l \quad \mu(l)[v_1] = l' \quad \mu(l') = w \quad \llbracket \odot \rrbracket (w, v_2) = w'}{\gamma \vdash^{\Gamma}_{stmt} & x[e_1] \odot = e_2 \ : \ \mu \rightleftharpoons \mu[l' \mapsto w']} \text{AssArrElemVar} \end{split}$$

$$\Gamma(c) = \left\langle \overbrace{(x_1, \dots, x_m)}^{fields}, methods \right\rangle \qquad \gamma(x) = l \qquad l_0 \notin \text{dom}(\mu) \dots l_m \notin \text{dom}(\mu)$$

$$\gamma \vdash_{stmt}^{\Gamma} \text{new } c \ x \ : \ \mu[l \mapsto 0] \rightleftharpoons \mu \left[l \mapsto l_0, l_0 \mapsto \left\langle c, (l_1, \dots, l_m) \right\rangle, l_1 \mapsto 0, \dots, l_m \mapsto 0 \right]$$
ObjNew

$$\frac{\langle l,\gamma\rangle \vdash^{\Gamma}_{stmt} \ \mathbf{new} \ c \ x \ : \ \mu' \rightleftharpoons \mu}{\gamma \vdash^{\Gamma}_{stmt} \ \mathbf{delete} \ c \ x \ : \ \mu \rightleftharpoons \mu'} \ \mathrm{ObjDelete}$$

$$\frac{\langle \gamma, \mu \rangle \vdash_{stmt}^{\Gamma} e \Rightarrow n \qquad \gamma(x) = l \qquad \mu(l) = 0 \qquad l' \notin \text{dom}(\mu)}{\gamma \vdash_{stmt}^{\Gamma} \text{new } a[e] \ x \ : \ \mu \rightleftharpoons \mu[l \mapsto l', l' \mapsto 0^n]} \text{ArrNew}$$

$$\frac{\langle l, \gamma \rangle \vdash_{stmt}^{\Gamma} \mathbf{new} \ a[e] \ x \ : \ \mu' \rightleftharpoons \mu}{\gamma \vdash_{stmt}^{\Gamma} \mathbf{delete} \ a[e] \ x \ : \ \mu \rightleftharpoons \mu'} \text{ARRDELETE}$$

Figure 1.25c: Extension to the semantic inference rules for statements in ROOPL++

$$\frac{\gamma(x) = l \qquad \gamma(x') = l' \qquad \mu(l) = v}{\gamma \vdash_{stmt}^{\Gamma} \mathbf{copy} \ c \ x \ x' : \ \mu[l' \mapsto 0] = \mu[l' \mapsto v]} \text{ Copy}$$

$$\frac{\langle l, \gamma \rangle \vdash_{stmt}^{\Gamma} \mathbf{copy} \ c \ x \ x' : \mu' \leftrightharpoons \mu}{\gamma \vdash_{stmt}^{\Gamma} \mathbf{uncopy} \ c \ x \ x' : \mu \rightleftharpoons \mu'} \mathbf{UNCOPY}$$

$$\begin{split} \langle \gamma, \mu \rangle \vdash^{\Gamma}_{stmt} \ e_1 \Rightarrow v_1 & \langle \gamma, \mu' \rangle \vdash^{\Gamma}_{stmt} \ e_2 \Rightarrow v_2 \\ \\ \frac{\mu' = \mu[l' \mapsto v_1, r \mapsto l']}{\gamma \vdash^{\Gamma}_{stmt} \ \mathbf{local} \ c \ x = e_1 \ s \ \mathbf{delocal} \ x = e_2 \ : \ \mu \rightleftharpoons \mu \end{split} \ \mathsf{LocalBlock} \end{split}$$

Figure 1.25d: Extension to the semantic inference rules for statements in ROOPL++ (cont)

all object fields, all of which, along with l_0 must be unused in μ . The result store μ'' is obtained after executing the body statement s in store μ' mapping x to object reference at l_0 , as long as all object fields are zero-cleared in μ'' afterwards. If any of these conditions fail, the object block statement is undefined.

Figures 1.25c and 1.25d show the extensions to the semantics of ROOPL with rules for **new/delete** and **copy/uncopy** statements, array element assignment and local blocks.

Rule ASSARRELEMVAR defines reversible assignment to array elements. After evaluating

expressions e_1 to v_1 and e_2 to v_2 , the value at the location of variable $x[v_1]$ under environment γ is mapped to the value v_3 resulting in an updated store μ' .

Dynamic object construction and destruction is defined by rules ObjNew and ObjDelete. For construction, x must be bound to a location l. We then make location l point to a new pair consisting of the class name and a vector of m new locations mapping object fields to locations. For destruction, x is still bound to l return l to a null pointer. As with object blocks, it is the program itself responsible for zero-clearing object fields before destruction. If the object fields are not zero-cleared, the ObjDelete statement is undefined.

Array construction and destruction is very similar to object construction and destruction. The major difference is we bind the location to a vector of size equal to the evaluated expression result. For deletion, we return the location of x to a null pointer and remove the binding to the vector from the store.

Object and array referencing is defined by rules COPY and UNCOPY. A reference is created and a new store μ' obtained by mapping x' to the reference r which x current maps to, if c matches the tuple mapped to the location l. A reference is removed and a new store μ' obtained if x and x' maps to the same reference r and x' then is removed from the store.

Local blocks are as previously mentioned, semantically similar to object blocks, where the memory location of variable x must be unused in the store μ . The updated store μ' contains location l' mapped to the evaluated value of e_1 , v_1 and the reference r, which x is bound to, mapped to l'. The result store after body statement execution, μ'' must have l' mapped to the expression value of e_2 , v_2 . Before the local block terminates, a third store update is executed, clearing the used memory locations, such that l' and r are mapped to zero and become unused again.

1.9.4 Programs

The judgment

$$\vdash_{prog} p \Rightarrow \sigma$$

defines the meaning of programs. The class p containing the main method is instantiated and the main function is executed with the partial function σ as the result, mapping variable identifiers to values, correlating to the class fields of the main class.

$$\Gamma = \operatorname{gen}(c_1, \dots, c_n) \qquad \overbrace{\Gamma(c) = \left(\{ \langle t_1, f_1 \rangle, \dots, \langle t_n, f_n \rangle \}, \text{ methods} \right)}$$

$$\left(\begin{array}{c} \left(\mathbf{method \ main} \ () \ s \right) \in \text{methods} \quad \gamma = [f_1 \mapsto 1, \dots, f_i \mapsto i] \\ \\ \mu = [1 \mapsto 0, \dots, i \mapsto 0, \ i+1 \mapsto \langle c, \gamma \rangle] \qquad \langle i+1, \gamma \rangle \vdash_{stmt}^{\Gamma} \ s \ : \ \mu \rightleftharpoons \mu' \\ \\ \vdash_{prog} \ c_1 \ \cdots \ c_n \Rightarrow (\mu' \circ \gamma) \end{array} \right)$$
MAIN

Figure 1.26: Semantic inference rules for programs, originally from [1]

As with ROOPL programs, the fields of the main method in the main class c are bound in a new environment, starting at memory address 1, as 0 is reserved for **nil**. The fields are zero-initialized in the new store μ and address i+1 which maps to the new instance of c. After body execution,

store μ' is obtained. The function $\mu' \circ \gamma$ maps class fields to their respective final values and serves as output of program p.

1.10 Program Inversion

In order to truly show that ROOPL++ in fact is a reversible language, we must demonstrate and prove local inversion of statements is possible, such that any program written in ROOPL++, regardless of context, can be executed in reverse. Haulund presented a statement inverter for ROOPL in [1], which maps statements to their inverse counterparts. Figure 1.27 shows the statement inverter, extended with the new ROOPL++ statements for construction/destruction and referencing copying/copy removal.

```
\mathcal{I}\llbracket s_1 \ s_2 \rrbracket = \mathcal{I}\llbracket s_2 \rrbracket \ \mathcal{I}\llbracket s_1 \rrbracket
\mathcal{I}[skip] = skip
                                                                                         \mathcal{I}[x -= e] = x += e
\mathcal{I}[\![x += e]\!] = x -= e
\mathcal{I}[x \triangleq e] = x \triangleq e
                                                                                         \mathcal{I}[x \iff e] = x \iff e
\mathcal{I}[x[e_1] += e_2] = x[e_1] -= e_2
                                                                                         \mathcal{I}[x[e_1] -= e_2] = x[e_1] += e_2
                                                                                         \mathcal{I}[x[e_1] \iff e_2] = x[e_1] \iff e_2
\mathcal{I}[x[e_1] \triangleq e_2] = x[e_1] \triangleq e_2
                                                                                         \mathcal{I}[\operatorname{copy} c \ x \ x'] = \operatorname{uncopy} c \ x \ x'
\mathcal{I}[\![\mathbf{new}\ c\ x]\!] = \mathbf{delete}\ c\ x
                                                                                         \mathcal{I}\llbracket \mathbf{uncopy} \ c \ x \ x' \rrbracket \ = \mathbf{copy} \ c \ x \ x'
\mathcal{I}[\![\mathbf{delete}\ c\ x]\!] = \mathbf{new}\ c\ x
\mathcal{I}[[\mathbf{call}\ q(\ldots)]] = \mathbf{uncall}\ q(\ldots)
                                                                                         \mathcal{I}[[\mathbf{call}\ x :: q(\ldots)]] = \mathbf{uncall}\ x :: q(\ldots)
\mathcal{I}[\mathbf{uncall}\ q(\ldots)] = \mathbf{call}\ q(\ldots)
                                                                                         \mathcal{I}[\mathbf{uncall}\ x :: q(\ldots)] = \mathbf{call}\ x :: q(\ldots)
\mathcal{I}[\mathbf{if}\ e_1\ \mathbf{then}\ s_1\ \mathbf{else}\ s_2\ \mathbf{fi}\ e_2]
                                                                                          = if e_1 then \mathcal{I}[s_1] else \mathcal{I}[s_2] fi e_2
\mathcal{I}[\mathbf{from}\ e_1\ \mathbf{do}\ s_1\ \mathbf{loop}\ s_2\ \mathbf{until}\ e_2]
                                                                                          = from e_1 do \mathcal{I}[s_1] loop \mathcal{I}[s_2] until e_2
\mathcal{I}[construct c \ x \ s \ destruct \ x]
                                                                                          = construct c \ x \ \mathcal{I}[s] destruct x
\mathcal{I}[\mathbf{local}\ t\ x\ = e \quad s \quad \mathbf{delocal}\ t\ x\ = e]
                                                                                          = local t x = e \mathcal{I}[s] delocal t x = e
```

Figure 1.27: ROOPL++ statement inverter, extended from [1]

Program inversion is conducted by recursive descent over components and statements. A proposed extension to the statement inverter for whole-program inversion, is retained in the ROOPL++ statement inverter. The extension covers the case, which reveals itself during method calling. As a method call is equivalent to an uncall with the inverse method and we simply change calls to uncalls during inversion, the inversion of the method body cancels out. The proposed extension, presented in [4, 1], simply avoids inversion of calls and uncalls, as shown in figure 1.28.

1.10.1 Invertibility of Statements

While the invertibility of statements remains untouched by the extensions made in ROOPL++, the following proof, originally presented in [1], has been included for completeness.