# Semantics of simple, object-oriented programming language (soopl)

#### Maciek Makowski

May 6, 2002

## 1 Abstract syntax

 $Actual Arguments ::= Expression \ list$ 

```
Program ::= (ClassMap, Initializer)
ClassMap ::= ClassDecl \ list
Initializer ::= Expression
ClassDecl ::= (Id, ClassBody)
Id ::= x_1, x_2, ...
ClassBody ::= (Superclass, ClassAttributes, ClassMethodMap, Attributes, MethodMap)
Superclass ::= Id
ClassAttributes ::= Attributes
ClassMethodMap ::= MethodMap
Attributes ::= Id \, list
MethodMap ::= MethodDecl \ list
MethodDecl ::= (Id, Arguments, MethodBody)
Arguments ::= Id \, list
MethodBody ::= Statement list
Statement ::= Assignment | Return | Evaluation | Execution
Assignment ::= (Id, Expression)
Return ::= Expression
Evaluation ::= Expression
Execution ::= Expression
Expression ::= Identifier|Constant|BlockDecl|Message|Creation
Identifier ::= Id
Constant ::= 0, 1, 2, ...
BlockDecl ::= Statement \, list
Message ::= (Target, Id, Actual Arguments)
Creation ::= Id
Target ::= Expression
```

#### 2 Semantic domains

- Identifiers:  $Id = \{x_1, x_2, ...\}$
- Locations:  $Loc = \{\alpha_1, \alpha_2, ...\}$
- Environments:  $Env = (Id \rightarrow DV_{\perp})$
- States:  $S = (Loc \rightarrow SV_{\perp})$
- Primitive values: Primitive = N
- Classes:  $Class = Loc \times Env \times Env \times Id \, list \times Env$ 
  - 1. superclass location
  - 2. class attributes environment
  - 3. class methods environment
  - 4. list of object attribute names
  - 5. object methods environment
- Objects:  $Object = Loc \times Env$ 
  - 1. class location
  - 2. object attributes environment
- Methods:  $Method = [Loc \rightarrow [Env \rightarrow [CExp \rightarrow CSt_{\perp}]]]$ 
  - 1. object/class location for which the method is being called
  - 2. local environment
  - 3. continuation to which the method result is passed
  - 4. final transformation on states
- Blocks:  $Block = Env \times Env \times Env \times (Env \rightarrow Env \rightarrow Env \rightarrow CSt \rightarrow CSt)$ 
  - 1. global environment
  - 2. attribute/variable environment
  - 3. method environment
  - 4. function, which receives these three environments and yields the statement continuation transformation
- Expression continuations:  $CExp = [EV \rightarrow CSt_{\perp}]$
- Statement continuation:  $CSt = [S \rightarrow S_{\perp} + \{error\}_{\perp}]$
- Storable values:  $SV = Class + Object + Primitive + Loc + \{empty\}$
- Denotable values:  $DV = Loc + CExp + Method \times Idlist + \{unbound\}$
- Expressable values: EV = Loc + Primitive

## 3 Miscellaneous operations

- ullet Obtaining the list's head: hd :  $lpha \ list 
  ightarrow lpha$
- Obtaining the list's tail: tl :  $\alpha list \rightarrow \alpha list$
- Concatenating two lists: append :  $\alpha \, list \rightarrow \alpha \, list \rightarrow \alpha \, list$
- Checking if element is in list: in list:  $\alpha \to \alpha \, list \to Boolean$
- Memory allocation: alloc :  $S \rightarrow Loc \times S$
- Creating the environment based on identifier list: makeenv :  $Id \, list \to S \to Env \times S$  makeenv =  $\lambda idl \in Id \, list. \lambda s \in S$ .

```
\begin{aligned} \text{if } idl &= [] \text{ then } (\lambda id \in Id. \text{unbound}, s) \\ \text{else let } (l,s') &= \text{alloc } s \text{ in} \\ \text{let } (\rho',s'') &= \text{makeenv } (\text{tl } idl) \, s' \text{ in} \\ (\rho'[l/\text{hd } idl],s'') \end{aligned}
```

• Narrowing the environment so that only identifiers bound in second environment are bound in the resulting environment: envcast :  $Env \rightarrow Env \rightarrow Env$ 

```
envcast = \lambda \rho_1 \in Env. \lambda \rho_2 \in Env.
\lambda id \in Id. if \rho_2 id = unbound then unbound else \rho_1 id
```

• Narrowing the environment so that only identifiers that are in the identifier list are bound in the resulting environment: envidcast :  $Env \rightarrow Id \, list \rightarrow Env$ 

```
\begin{array}{l} \mathsf{envidcast} \, = \lambda \rho \in Env. \\ \lambda id \in Id. \ \mathsf{if} \ \mathsf{inlist} \ id \ idl \ \mathsf{then} \ \rho \, id \\ \mathsf{else} \ \mathsf{unbound} \end{array}
```

• Composition of two environments: envcomb :  $Env \rightarrow Env \rightarrow Env$  envcomb =  $\lambda \rho_1 \in Env. \lambda \rho_2 \in Env.$  $\lambda id \in Id.$  if  $\rho_2 id$  = unbound then  $\rho_1 id$ 

$$\lambda id \in Id$$
. if  $\rho_2 id = \text{unbound then } \rho_1 id$  else  $\rho_2 id$ 

## 4 Additional assumptions

In the initial environment identifier Root is bound to location  $\alpha_r$ . In the initial state in location  $\alpha_r$  there is a class:  $(\alpha_r, \lambda id \in Id.$ unbound,  $\lambda id \in Id.$ unbound,  $[], \lambda id \in Id.$ unbound). This is the root class of the whole inheritance tree.

#### **5** Semantic functions

$$\begin{split} \mathcal{M} Program : Program \rightarrow Env \rightarrow S \rightarrow Env \times S \\ \mathcal{M} ClassMap : ClassMap \rightarrow Env \rightarrow S \rightarrow Env \times S \\ \mathcal{M} Initializer : Initializer \rightarrow Env \rightarrow S \rightarrow S \\ \mathcal{M} ClassDecl : ClassDecl \rightarrow Env \rightarrow S \rightarrow Env \times S \end{split}$$

```
 \begin{split} \mathcal{M}ClassBody: ClassBody \rightarrow Env \rightarrow S \rightarrow Loc \times S \\ \mathcal{M}ClMethodMap: MethodMap \rightarrow Env \rightarrow Env \rightarrow Env \rightarrow S \rightarrow Env \\ \mathcal{M}ClMethodDecl: MethodDecl \rightarrow Env \rightarrow Env \rightarrow Env \rightarrow S \rightarrow Env \\ \mathcal{M}MethodMap: MethodMap \rightarrow Env \rightarrow Env \rightarrow Env \rightarrow Id\, list \rightarrow Env \rightarrow S \rightarrow Env \\ \mathcal{M}MethodDecl: MethodDecl \rightarrow Env \rightarrow Env \rightarrow Env \rightarrow Id\, list \rightarrow Env \rightarrow S \rightarrow Env \\ \mathcal{M}MethodDecl: MethodBody \rightarrow Env \rightarrow Env \rightarrow Env \rightarrow CSt \\ \mathcal{M}Statement: Statement \rightarrow Env \rightarrow Env \rightarrow Env \rightarrow CSt \\ \mathcal{M}StatementList: Statement\, list \rightarrow Env \rightarrow Env \rightarrow Env \rightarrow CSt \rightarrow CSt \\ \mathcal{M}Expression: Expression \rightarrow Env \rightarrow Env \rightarrow Env \rightarrow Env \rightarrow Id\, list \rightarrow S \rightarrow S \\ \mathcal{M}Arguments: ActualArguments \rightarrow Env \rightarrow Env \rightarrow Env \rightarrow Env \rightarrow Id\, list \rightarrow S \rightarrow S \\ \end{split}
```

### **6** Semantic equations

**Program:** process the class map and then evaluate the initializer in resulting environment and state.

```
\mathcal{M}Program[[mkProgram(clm,init)]] = \lambda \rho \in Env. \lambda s \in S. let (\rho',s') = \mathcal{M}ClassMap[[clm]] \rho s in (\rho',\mathcal{M}Initializer[[init]] \rho' s')
```

**Class map:** process all the class declarations in the list, each one in environment/state resulting from previous declaration.

```
\begin{split} \mathcal{M}ClassMap[[mkClassMap([])]] &= \lambda \rho \in Env. \lambda s \in S.(\rho,s) \\ \mathcal{M}ClassMap[[mkClassMap(cdl)]] &= \lambda \rho \in Env. \lambda s \in S. \\ &\text{let } (\rho',s') = \mathcal{M}ClassDecl[[hd\ cdl]]\ \rho\ s\ \text{in } \mathcal{M}ClassMap[[mkClassMap(tl\ cdl)]]\ \rho'\ s' \end{split}
```

**Initializer:** treat the initializer as evaluation of the expression where attribute and method environments are empty and the statement continuation is identity.

```
\mathcal{M}Initializer[mkInitializer(exp)]] = \lambda \rho \in Env. \lambda s \in S.
\mathcal{M}Statement[mkEvaluation(exp)]] \rho
(\lambda id \in Id.unbound) (\lambda id \in Id.unbound) (\lambda s \in S.s) s
```

**Class declaration:** process the class body and bind the identifier to the location where the class is placed.

```
\mathcal{M}ClassDecl[mkClassDecl(id, cb)] = \lambda \rho \in Env.\lambda s \in S.
let (l, s') = \mathcal{M}ClassBody[cb][\rho s \text{ in } (\rho[l/id], s')
```

**Class body:** find the superclass for the class being processed and:

- create the class attribute environment and combine the superclass' attribute environment with it
- process the class method map

- append the object attribute identifier list to the superclass' object attribute identifier list
- process the object method map

after that, allocate memory for the class and put the class in the allocated location.

```
\begin{split} \mathcal{M}ClassBody[[mkClassBody(id, catr, cmm, atr, mm)]] &= \lambda \rho \in Env. \lambda s \in S. \\ \text{let } mkClass(p, \rho_{ca}, \rho_{cm}, al, \rho_m) &= s\left(\rho \, id\right) \text{ in} \\ \text{let } (\rho'_{ca}, s') &= \text{makeenv } catr \, s \text{ in} \\ \text{let } \rho''_{ca} &= \text{envcomb } \rho_{ca} \, \rho'_{ca} \text{ in} \\ \text{let } \rho'_{ca} &= \text{envCl} Method Map[[cmm]] \, \rho \, \rho''_{ca} \, \rho_{cm} \, s' \\ al' &= \text{append } atr \, al \text{ in} \\ \text{let } \rho'_{m} &= \mathcal{M} Method Map[[mm]] \, \rho \, \rho''_{ca} \, \rho'_{cm} \, al' \, \rho_{m} \, s' \\ (l, s'') &= \text{alloc } s' \text{ in} \\ (l, s''[mkClass(\rho \, id, \rho''_{ca}, \rho'_{cm}, al', \rho'_{m})/l]) \end{split}
```

**Class method map:** process the class method declarations in the list, each one using class method environment resulting from previous declarations.

```
\mathcal{M}ClMethodMap[\llbracket mkMethodMap([])\rrbracket] = \lambda\rho \in Env. \\ \lambda\rho_{ca} \in Env. \\ \lambda\rho_{c
```

Class method declaration: construct the operator F that takes method m as an argument, processes method body combining narrowed class attribute environment with local method environment (containing method arguments), narrowing the class method environment and binding the id of declared method to m in this environment. Then compute the fixpoint of F and bind the id of declared method to this fixpoint. Note the narrowing, which ensures, that only the identifiers bound in the environments of the class in which the method is declared are accessible to the method.

```
\begin{split} \mathcal{M}ClMethodDecl[[mkMethodDecl(id,args,mb)]] &= \lambda \rho \in Env. \lambda \rho_{ca} \in Env. \lambda \rho_{cm} \in Env. \\ \lambda s \in S. \\ &\text{let } F = \lambda m \in Method. \lambda t \in Loc. \lambda \rho_{loc} \in Env. \lambda k \in CExpr. \lambda s \in S. \\ &\text{let } mkClass(p, \rho'_{ca}, \rho'_{cm}, al, \rho_m) = st \text{ in } \\ &\mathcal{M}MethodBody[[mb]] \rho \\ &\text{ (envcomb (envcast } \rho'_{ca} \rho_{ca}) \rho_{loc}[k/\text{return}][t/\text{self}]) \\ &\text{ (envcast } \rho'_{cm} \rho_{cm})[m/id] s \\ &\text{ in } \rho_{cm}[(fix F, args)/id] \end{split}
```

**Object method map:** process the object method declarations in the list, each one using object method environment resulting from previous declarations.

```
\mathcal{M}MethodMap[[mkMethodMap([])]] = \lambda \rho \in Env. \lambda \rho_{ca} \in Env. \lambda \rho_{cm} \in Env. \lambda al \in Id \, list.\lambda \rho_{m} \in Env. \lambda s \in S. \rho_{m}\mathcal{M}MethodMap[[mkMethodMap(mdl)]] = \lambda \rho \in Env. \lambda \rho_{ca} \in Env. \lambda \rho_{cm} \in Env.
```

```
\lambda al \in Id\, list. \\ \lambda \rho_m \in Env. \\ \lambda s \in S. \\ \mathcal{M}MethodMap[\![mkMethodMap(\mathsf{tl}\, mdl)]\!] \rho \, \rho_{ca} \, \rho_{cm} \, al} \\ (\mathcal{M}MethodDecl[\![\mathsf{hd}\, mdl]\!] \, \rho \, \rho_{ca} \, \rho_{cm} \, al \, \rho_m \, s) \, s
```

**Object method declaration:** construct the operator F that takes method m as an argument, processes method body combining narrowed object attribute environment with local method environment (containing method arguments), narrowing the object method environment and binding the id of declared method to m in this environment. Then compute the fixpoint of F and bind the id of declared method to this fixpoint. Narrowing works the same way as with class method declaration.

```
\begin{split} \mathcal{M}MethodDecl[[mkMethodDecl(id, args, mb)]] &= \lambda \rho \in Env. \lambda \rho_{ca} \in Env. \lambda \rho_{cm} \in Env. \\ \lambda al \in Id\, list. \lambda \rho_m \in Env. \lambda s \in S. \\ &\text{ let } F = \lambda m \in Method. \lambda t \in Loc. \lambda \rho_{loc} \in Env. \lambda k \in CExp. \lambda s \in S. \\ &\text{ let } mkObject(c, \rho_a) = s\, t \text{ in } \\ &\text{ let } mkClass(p, \rho'_{ca}, \rho'_{cm}, al', \rho'_m) = s\, c \text{ in } \\ & \qquad \qquad \mathcal{M}MethodBody[\![mb]\!] \, \rho \\ & \qquad \qquad \qquad (\text{envcomb (envidcast } \rho_a\, al) \, \rho_{loc}[k/\text{return}][t/\text{self}]) \\ & \qquad \qquad \qquad (\text{envcast } \rho'_m \, \rho_m)[m/id]\, s \end{split} in \rho_m[(fix\, F, args)/id]
```

**Method body:** process the statements in the list, each one with statement continuation resulting from *following* statements; at the end of the list put the continuation that takes the expression continuation bound to identifier return in the attribute environment and passes to it the location bound to self in this environment. Thanks to this, each method that doesn't explicitly return something, will return the class/object it was called for (like in Smalltalk).

```
\mathcal{M}MethodBody[[mkMethodBody([])]] = \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env.\lambda s \in S.(\rho_a \text{ return}) (\rho_a \text{ self}) s\mathcal{M}MethodBody[[mkMethodBody(stl)]] = \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env.\mathcal{M}Statement[[hd stl]] \rho \rho_a \rho_m (\mathcal{M}MethodBody[[mkMethodBody(tl stl)]] \rho \rho_a \rho_m)
```

**Assignment:** evaluate the expression with expression continuation that puts given value under location bound to id in the attribute environment.

```
\mathcal{M}Statement[[mkAssignment(id, exp)]] = \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env. \lambda k \in CSt.
\mathcal{M}Expression[[exp]] \rho \rho_a \rho_m (\lambda v \in EV. \lambda s \in S.k s[v/\rho_a id])
```

**Return:** evaluate the expression with expression continuation that passes given value to the expression continuation bound to return in the attribute environment.

```
\mathcal{M}Statement[[mkReturn(exp)]] = \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env. \lambda k \in CSt.
\mathcal{M}Expression[[exp]] \rho \rho_a \rho_m (\lambda v \in EV. \lambda s \in S. \rho_a \text{ return } v s)
```

**Evaluation:** evaluate the expression with expression continuation that discards the given value (only the side-effects will count).

```
\mathcal{M}Statement[[mkEvaluation(exp)]] = \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env. \lambda k \in CSt.
\mathcal{M}Expression[[exp]] \rho \rho_a \rho_m (\lambda v \in EV.k)
```

**Execution:** evaluate the expression with expression continuation that receives a location of a block and then executes the block in saved environments.

```
\mathcal{M}Statement[\![mkExecution(exp)]\!] = \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env. \lambda k \in CSt.
\mathcal{M}Expression[\![exp]\!] \rho \rho_a \rho_m (\lambda v \in EV. \lambda s \in S.
\operatorname{let} mkBlock(\rho, \rho_a, \rho_m, bl) = s v \text{ in}
bl \rho \rho_a \rho_m k s)
```

**Statement list:** process all the statements in the list, each with statement continuation resulting from *following* statements (very similar to method body).

```
\mathcal{M}StatementList[[]] = \lambda \rho \in Env. \\ \lambda \rho_a \in Env. \\ \lambda \rho_m \in Env. \\ \lambda \rho_m \in Env. \\ \lambda k \in CSt. \\ \mathcal{M}Statement[[stl]] = \lambda \rho \in Env. \\ \lambda \rho_a \in Env. \\ \lambda \rho_m \in Env. \\ \lambda k \in CSt. \\ \mathcal{M}Statement[[tl]] \rho \rho_a \rho_m (\mathcal{M}StatementList[[tl]] \rho \rho_a \rho_m k)
```

**Identifier:** if given id is unbound in the attribute environment then get the value bound to this id in the global environment, otherwise get the value bound to it in the attribute environment; pass the value to given expression continuation.

```
\mathcal{M}Expression[[mkIdentifier(id)]] = \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env. \lambda k \in CExp.
\lambda s \in S. if \rho_a id = \text{unbound then } k (\rho id) s \text{ else } k (s (\rho_a id)) s
```

**Constant:** pass the constant value to given expression continuation.

```
\mathcal{M}Expression[mkConstant(n)] = \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env. \lambda k \in CExp. k n
```

**Block declaration:** allocate memory for the block, process the statement list that constitutes the block and put the result in allocated location together with current environments; pass the block location to given expression continuation.

```
\begin{split} \mathcal{M}Expression[\![mkBlockDecl(stl)]\!] &= \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env. \lambda k \in CExp. \\ \lambda s \in S. \text{let } (l,s') &= \text{alloc } s \text{ in} \\ &\quad k \, l \, s'[\![mkBlock(\rho,\rho_a,\rho_m,\lambda\rho \in Env. \lambda\rho_a \in Env. \lambda\rho_m \in Env. \lambda k \in CSt. \\ &\quad \mathcal{M}StatementList[\![stl]\!] \, \rho \, \rho_a \, \rho_m \, k)/l] \end{split}
```

**Message:** first evaluate the *target* expression; then:

- determine if the target is object or class
- find the method using target's method environment

- create local variable environment from the method argument identifier list
- combine the target attribute environment with it
- process the arguments

finally, call the method passing target location, local environment, given expression continuation and the state resulting from processing method arguments.

```
\mathcal{M}Expression[mkMessage(exp,id,args)] = \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env.
             \lambda k \in CExp.
             \mathcal{M}Expression[\![exp]\!] \rho \rho_a \rho_m (\lambda t \in Loc.\lambda s \in S.
                             if st \in Object then
                                         \mathrm{let}\ mkObject(c,\rho_a') = s\,t\ \mathrm{in}
                                                       let mkClass(p, \rho_{ca}, \rho_{cm}, al, \rho'_{m}) = s c in
                                                                     let (m, ids) = \rho'_m id in
                                                                                   let (\rho_{loc}, s') = \text{makeenv } ids s \text{ in}
                                                                                                let \rho_{loc} = \text{envcomb } \rho'_a \, \rho_{loc} in
                                                                                                              \mathsf{let}\ s'' = \mathcal{M}Arguments[\![args]\!]
                                                                                                                            else
                                         let mkClass(p, \rho_{ca}, \rho_{cm}, al, \rho'_{m}) = st in
                                                       \mathsf{let}\;(m,ids) = \rho_{cm}\,id\;\mathsf{in}
                                                                     let (\rho_{loc}, s') = makeenv ids s in
                                                                                  \begin{array}{c} \text{let } \rho_{loc} = \text{envcomb } \rho_a' \, \rho_{loc} \text{ in} \\ \text{let } s'' = \mathcal{M}Arguments[\![args]\!] \end{array}

ho 
ho_a 
ho_m 
ho_{loc} ids \, s' in m \, t \, 
ho_{loc} \, k \, s'')
```

**Creation:** find the class bound to id, create the object attribute environment using the id list in the class, allocate memory for the object and put a new object into resulting location; pass the object location to given expression continuation.

```
 \begin{split} \mathcal{M}Expression[\![mkCreation(id)]\!] &= \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env. \lambda k \in CExp. \\ \lambda s \in S. \text{let } (l,s') &= \text{alloc } s \text{ in} \\ &\text{let } mkClass(p,\rho_{ca},\rho_{cm},al,\rho'_m) = s'\left(\rho\,id\right) \text{ in} \\ &\text{let } (\rho_a,s'') = \text{makeenv } al\,s' \text{ in} \\ &\qquad \qquad k\,l\,s''[mkObject(\rho\,id,\rho_a)/l] \end{split}
```

**Arguments:** evaluate each expression in the expression list with the expression continuation that puts given value under the location bound to the next identifier in the argument id list. If there are more expressions than argument identifiers, don't evaluate the surplus expressions.

```
\begin{split} \mathcal{M} Arguments[\![mkArguments([])]\!] &= \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env. \lambda \rho_{loc} \in Env. \\ \lambda ids \in Id\, list. \lambda s \in S.s \\ \mathcal{M} Arguments[\![mkArguments(expl)]\!] &= \lambda \rho \in Env. \lambda \rho_a \in Env. \lambda \rho_m \in Env. \lambda \rho_{loc} \in Env. \\ \lambda ids \in Id\, list. \end{split}
```

$$\begin{split} \text{if } ids &= [] \text{ then } \lambda s \in S.s \\ \text{else let } k &= \mathcal{M}Arguments[\![mkArguments(\mathsf{tl}\,expl)]\!] \ \rho \ \rho_a \ \rho_m \ \rho_{loc} \ (\mathsf{tl}\,ids) \ \text{in} \\ \mathcal{M}Expression[\![\mathsf{hd}\,expl]\!] \ \rho \ \rho_a \ \rho_m \ (\lambda v \in EV.\lambda s \in S.k \ s[v/\rho_{loc} \ (\mathsf{hd}\,ids)]) \end{split}$$