

# MITScript Language Specification

## 1 Preliminaries

We denote MITScript expressions by  $e \in E$  and MITScript statements by  $s \in S$ . A program variable or field name  $x \in X$  is an element of the set of all valid program variables and field names  $X$ .

**Values.** A value  $v \in V$  in the program can be one of the following six categories: Bool, Integer, String, Record, Function, or None. We denote a boolean value by the constructor  $\text{Bool}(b)$  where  $b \in \{\text{True}, \text{False}\}$ . We denote an integer value by the constructor  $\text{Integer}(n)$  where  $n \in Z$ . We denote a string value by the constructor  $\text{String}(st)$  where  $st$  is a string in the language of valid MITScript strings. We denote a record by the constructor  $\text{Record}(m)$  where  $m \in M = X \rightarrow A$  is a *map*, mapping a program variable or field name to an address  $a \in A = (N \cup \cdot)$ . An address is a natural number or the distinguished value  $\cdot$  which represents an invalid address (e.g., similar to NULL in other languages).

We denote a function value by the constructor  $\text{Function}(a, x, s)$ , where  $a$  is an address of the function's *frame* (as described below),  $x$  is the function's formal argument, and  $s$  is the MITScript code corresponding to the function's body. The value None is a special singleton value, similar to null.

In your implementation you should expect to use objects to represent all the different types of values. Specifically, it is recommended that you define a Value class, and then make each of the different value types subclasses of Value.

**Program State.** We represent the state of an MITScript program by a stack and a heap. Because of closures, however, we model the stack itself as a series of addresses of heap allocated stack frames.

A frame  $\sigma \in \Sigma = \{X \cup \{\text{parent}, \text{global}\}\}A$  is a mapping from program variables and field names to addresses that store in the heap the contents of the program variables. A frame has two additional distinguished fields. The *parent* field stores a pointer to the current frame's parent frame. The *global* field stores a pair  $(X_g, a)$ , where  $X_g$  is a set of program variables that have been declared as global within the frame (by `global` statements) and  $a$  the address of the global frame.

The state of the program includes a heap  $h : A \rightarrow v$  that maps addresses  $a$  to heap allocated values  $v$  which can either be stack frames or one of the six types of values listed above.

It also includes a stack  $\gamma$  of references to stack frames. We will use the notation  $\gamma = \gamma'; a$  to indicate that the stack  $\gamma$  is equal to another stack  $\gamma'$  with an address  $a$  added at the top. Similarly, we use the notation  $\gamma = a_1; \gamma'; a_2$  to indicate that  $\gamma$  is a stack that has  $a_1$  at the bottom, followed by another stack  $\gamma'$  followed by an address  $a_2$  at the top.

**Native Functions.** Your implementation should support the three native, predefined global functions described in Section 5,

**Execution.** The program starts its execution with a stackframe  $\sigma = \{\text{parent} : \cdot\}$  with a parent pointer equal to  $\cdot$ . The initial heap contains only this initial global stackframe and a pre-allocated value for `None`, ( $h_0 = \{a_g : \sigma, a_{\text{None}} : \text{None}\}$ ). The initial stack contains just the global stack frame in the heap  $\gamma_0 = a_g$  as well as the definitions that you need to appropriately support MITScript's native functions.

**Errors.** When a program performs an illegal operation, execution should stop, and your interpreter must report one of the following exceptions:

- **UninitializedVariableException:** when the program attempts to read from a variable in the stack-frame that has not been initialized.
- **IllegalCastException:** when the program attempts to apply an operation to value type for which it does not apply.
- **IllegalArithmeticException:** when the program attempts to divide by zero.
- **RuntimeException:** other error situations, such as passing too many arguments to a function.

In the semantics below we indicate exactly which of these three exceptions your interpreter should throw.

## 2 Statements

Figure 1 presents the semantics of statements. The evaluation relation for statements is of the form  $(\gamma, h, s) \rightarrow \eta$  which denotes that execution of a statement  $s$  from a stack  $\gamma$  and a heap  $h$  evaluates to a *heap-value*  $\eta$ . A heap value  $\eta$  is either a heap  $h \in H$  or a *return-value* denoted by the constructor  $\text{Return}(h, v)$  where  $h$  is a heap and  $v$  is a value. A distinguished return-value enables the semantics specification to appropriately handle cases where `return` statements appear arbitrarily within the body of a function.

<b>VARASSIGNMENT</b> $\frac{(\gamma, h, e) \rightarrow (h', v) \quad \text{lookup\_write}(a_1, h', x) = a_2 \quad \sigma = h'(a_2) \quad a_3 \notin \text{dom}(h) \quad \sigma' = \sigma[x : a_3]}{(\gamma; a_1, h, x = e) \rightarrow h'[a_3 : v][a_2 : \sigma]}$		
<b>HEAPASSIGNMENT</b> $\frac{(\gamma, h, e_1) \rightarrow (h', \text{Record}(a_1)) \quad (\gamma, h', e_2) \rightarrow (h'', v) \quad h''(a_1) = m \quad a_2 \notin \text{dom}(h'') \quad m' = m[x : a_2]}{(\gamma, h, e_1.x = e_2) \rightarrow h''[a_2 : v][a_1 : m']}$		
<b>HEAPINDEXASSIGNMENT</b> $\frac{x = \text{toString}(v_1) \quad (\gamma, h, e_1) \rightarrow (h', \text{Record}(a_1)) \quad (\gamma, h', e_2) \rightarrow (h'', v_1) \quad (\gamma, h'', e_3) \rightarrow (h''', v_2) \quad h'''(a) = m \quad a_2 \notin \text{dom}(h'') \quad m' = m[x : a_2]}{(\gamma, h, e_1[e_2] = e_3) \rightarrow h'''[a_2 : v_2][a_1 : m']}$		
<b>IFTURE</b> $\frac{(\gamma, h, e) \rightarrow (h', \text{Bool}(\text{True})) \quad (\gamma, h', s_1) \rightarrow \eta}{(\gamma, h, \text{if } e s_1 \text{ else } s_2) \rightarrow \eta}$	<b>IFFALSE</b> $\frac{(\gamma, h, e) \rightarrow (h', \text{Bool}(\text{False})) \quad (\gamma, h', s_2) \rightarrow \eta}{(\gamma, h, \text{if } e s_1 \text{ else } s_2) \rightarrow \eta}$	
<b>WHILE</b> $\frac{(\gamma, h, \text{if } e (s ; \text{while } e s) \text{ else } \epsilon) \rightarrow \eta}{(\gamma, h, \text{while } e s) \rightarrow \eta}$	<b>SEQUENCE</b> $\frac{(\gamma, h, s_1) \rightarrow h' \quad (\gamma, h', s_2) \rightarrow \eta}{(\gamma, h, s_1 ; s_2) \rightarrow \eta}$	
<b>SEQUENCERETURN</b> $\frac{(\gamma, h, s_1) \rightarrow \text{Return}(h', v)}{(\gamma, h, s_1 ; s_2) \rightarrow \text{Return}(h', v)}$	<b>GLOBAL</b> $\frac{}{(\gamma, h, \text{global } x) \rightarrow h}$	<b>RETURN</b> $\frac{(\gamma, h, e) \rightarrow (h', v)}{(\gamma, h, \text{return } e) \rightarrow \text{Return}(h', v)}$

Figure 1: Semantics of statements

**VarAssignment.** The VarAssignment rule evaluates an expression and assigns the resulting value to the address of the variable in the heap. The `lookup_write` function (defined below) returns the address of the frame in which to update  $x$  using the appropriate scoping rules.

**HeapAssignment.** In this case, the assignment is not updating a stack frame; it is updating a record in the heap. The statement should throw an `IllegalCastException` if the result of the evaluation of  $e_1$  is not a record.

**HeapIndexAssignment.** This rule is similar to the previous one, except the field name is not hardcoded; instead, the index expression needs to be evaluated and cast to the string according to the `toString` function defined below. The operation should throw an `IllegalCastException` if the result of evaluation  $e_1$  is not a record.

**If.** If the condition evaluates to true, the result of the if statement is the result of evaluating the then branch. There is a symmetric case for when the branch condition evaluates to false. The statement should throw an `IllegalCastException` if the result of the evaluation of  $e$  is not a boolean.

**While.** The statement should throw an `IllegalCastException` if the result of  $e$  is not a boolean.

**Sequence.** The operator evaluates the first statement and then evaluates the second statement. Note that if the first statement yields a return-value, then the program skips the execution of the second statement.

**Global.** The statement is a no-op; its semantics is only incorporated later during function calls.

**Return.** The statement evaluates the expression and yields a return-value.

**Lookup.** The `lookup_write` function identifies the address of the appropriate frame into which to write. There are two cases to consider:

- $x \in g$ , so  $x$  is a global variable. In that case, we need to update the value of  $x$  in the global stack frame (the `GlobalLookup` rule).
- $x \notin g$ , so  $x$  is a local variable. In that case, we need to update the value of  $x$  in the local stack frame. (the `LocalLookup` rule)

$\frac{\text{GLOBALLOOKUP} \quad h(a_1) = \sigma_1 \quad \sigma_1(\text{global}) = (X_g, a_2) \quad x \in X_g}{\text{lookup\_write}(a_1, h, x) = a_2}$	$\frac{\text{LOCALLOOKUP} \quad h(a_1) = \sigma_1 \quad \sigma_1(\text{global}) = (X_g, a_2) \quad x \notin X_g}{\text{lookup\_write}(a_1, h, x) = a_1}$
--	--

Figure 2: Semantics of the `lookup_write` function

### 3 Expressions

Figures 3, 4, and 5 present the semantics of expressions. The evaluation relation for expressions is of the form  $(\gamma, h, e) \rightarrow (h', v)$  which denotes that given a stack  $\gamma$  and a heap  $h$ , an expression  $e$  evaluates to a new heap  $h'$  and a value  $v$ .

#### 3.1 Constants

Figure 3 presents the semantics of constants. For each type of constant value, the interpreter generates a value corresponding to the constant.

INTEGERCONSTANT	BOOLEANCONSTANTTRUE	BOOLEANCONSTANTFALSE
$(\gamma, h, n) \rightarrow (h, \text{Integer}(n))$	$(\gamma, h, \text{true}) \rightarrow (h, \text{Bool}(\text{True}))$	$(\gamma, h, \text{false}) \rightarrow (h, \text{Bool}(\text{False}))$
STRINGCONSTANT		NONECONSTANT
$(\gamma, h, st) \rightarrow (h, \text{String}(st))$		$(\gamma, h, \text{None}) \rightarrow (h, \text{None})$

Figure 3: Constant Expressions

### 3.2 Arithmetic and Logical Expressions

Figure 4 presents the semantics of arithmetic and logical expressions. We partition the operators into several categories: logical operations,  $lop \in Lop \rightarrow \& \mid |$ , arithmetic operations,  $op \in Op \rightarrow + \mid - \mid \times \mid /$ , unary operations,  $-$  and  $!$ , comparison operations,  $cmp \in Cmp \rightarrow < \mid > \mid <= \mid >=$ , and then equality ( $==$ ). We also assume the existence of a function eval that can mathematically evaluate a given operation and its arguments.

**Logical Operations.** Logical operators can only be applied to boolean values.

**Arithmetic Operations.** The arithmetic operators ‘-’ ‘\*’ and ‘/’ can only be applied to integer values. The operator ‘+’ on two integers corresponds to standard integer addition. The operator ‘+’ on two strings corresponds to string concatenation. The operator ‘+’ on a string and a value that is not a string causes the non-string value to be cast to a string.

**Unary Operations.** The unary minus operator ‘-’ can only be applied to an integer value. The unary not operator ‘!’ can only be applied to a boolean value.

**Comparison Operations.** Comparison operators  $<, >, <=, >=$  are only defined for integers.

**Equality Operations.** Equality between two integers, two booleans, two strings, or between two None values performs a value comparison. Equality between two Records checks for pointer equality. Equality between two Functions checks that you have the same frame address and the same function. Equality between two different types of values returns false.

**Errors.** Your interpreter should generate an `IllegalCastException` for operations that attempt to on types that are not given an explicit semantics. Your interpreter should generate an `IllegalAritheticException` for division by zero.

### 3.3 Data and Function Expressions

Figure 5 presents the semantics of data access expressions as well as function expressions.

**Variable Read.** VariableRead rule specifies the semantics of reading from a variable. The `lookup_read` function (defined below) returns the address of  $x$  using the appropriate lookup results. If the variable has not been defined in any available stack frame, then your interpreter should generate an `UninitializedVariableException`.

**Record Constructor** There record constructor just evaluates all its initializer expressions and constructs a Record value (a map) from fields to values. Note that the initializer expressions must be evaluated in order.

$\text{LOGICALOPERATION}$ $\frac{(\gamma, h, e_1) \rightarrow (h', \text{Bool}(b_1)) \quad (\gamma, h', e_2) \rightarrow (h'', \text{Bool}(b_2))}{(\gamma, h, e_1 \text{ lop } e_2) \rightarrow (h'', \text{Bool}(\text{eval}(\text{lop}, b_1, b_2)))}$	$\text{ARITHMETICOPERATION}$ $\frac{(\gamma, h, e_1) \rightarrow (h', \text{Integer}(n_1)) \quad (\gamma, h', e_2) \rightarrow (h'', \text{Integer}(n_2))}{(\gamma, h, e_1 \text{ op } e_2) \rightarrow (h'', \text{Integer}(\text{eval}(\text{op}, n_1, n_2)))}$	$\text{UNARYMINUS}$ $\frac{(\gamma, h, e) \rightarrow (h', \text{Integer}(n))}{(\gamma, h, -e) \rightarrow (h', \text{Integer}(-n))}$
$\text{UNARYNOT}$ $\frac{(\gamma, h, e) \rightarrow (h', b)}{(\gamma, h, !e) \rightarrow (h', \text{Bool}(\neg b))}$	$\text{COMPARISONOPERATION}$ $\frac{(\gamma, h, e_1) \rightarrow (h', \text{Integer}(n_1)) \quad (\gamma, h', e_2) \rightarrow (h'', \text{Integer}(n_2))}{(\gamma, h, e_1 \text{ cmp } e_2) \rightarrow (h'', \text{Bool}(\text{eval}(\text{cmp}, n_1, n_2)))}$	
$\text{STRINGCONCATENTATION}$ $\frac{(\gamma, h, e_1) \rightarrow (h', \text{String}(st_1)) \quad (\gamma, h', e_2) \rightarrow (h'', v) \quad st_2 = \text{toString}(v)}{(\gamma, h, e_1 + e_2) \rightarrow (h'', \text{String}(st_1 . st_2))}$	$\text{NONEQUALITY}$ $\frac{(\gamma, h, e_1) \rightarrow (h', \text{None}) \quad (\gamma, h', e_2) \rightarrow (h'', \text{None})}{(\gamma, h, e_1 == e_2) \rightarrow (h'', \text{Bool}(\text{True}))}$	
$\text{PRIMITIVEEQUALITY}$ $\frac{(\gamma, h, e_1) \rightarrow (h', t_1(v_1)) \quad (\gamma, h', e_2) \rightarrow (h'', t_2(v_2)) \quad t_1 = t_2 \quad \{t_1, t_2\} \in \{\text{Integer}, \text{Bool}, \text{String}\}}{(\gamma, h, e_1 == e_2) \rightarrow (h'', \text{Bool}(v_1 == v_2))}$	$\text{RECORDEQUALITY}$ $\frac{(\gamma, h, e_1) \rightarrow (h', \text{Record}(a_1)) \quad (\gamma, h', e_2) \rightarrow (h'', \text{Record}(a_2))}{(\gamma, h, e_1 == e_2) \rightarrow (h'', \text{Bool}(a_1 == a_2))}$	
$\text{FUNCTIONEQUALITYTRUE}$ $\frac{(\gamma, h, e_1) \rightarrow (h', \text{Function}(a_1, x, s)) \quad (\gamma, h', e_2) \rightarrow (h'', \text{Function}(a_2, x, s))}{(\gamma, h, e_1 == e_2) \rightarrow (h'', \text{Bool}(a_1 == a_2 \wedge x_1 == x_2 \wedge s_1 == s_2))}$	$\text{PRIMITIVEEQUALITYMISMATCHED}$ $\frac{(\gamma, h, e_1) \rightarrow (h', t_1(v_1)) \quad (\gamma, h', e_2) \rightarrow (h'', t_2(v_2)) \quad t_1 \neq t_2}{(\gamma, h, e_1 == e_2) \rightarrow (h'', \text{Bool}(\text{False}))}$	

Figure 4: Arithmetic and Logical Expressions

**Field Read.** Evaluating the expression  $e$  produces a Record; a pointer to a map from variables to values, and the variable ' $x$ ' is mapped to the value  $v$ . If the map does not contain a field  $x$ , the value should be `None`. The statement should throw an `IllegalCastException` if the result of the evaluation of  $e$  is not a record.

**Index Read.** When the index evaluates to a string, that string is used as a variable name to index into the map produced by evaluating  $e_1$ . If  $v_1$  is not a string, then it should be converted to a string by the casting rules below. Like before, if there is no field  $x$  in  $h'(a)$ , then Index Read should return `None`.

**Function Creation.** The construction of a function creates a new function value that captures the current frame.

**Function Call** There are multiple steps to this rule. First, the base expression for the function call must be evaluated, and it must evaluate to a function value, which consists of the address of a frame, and the code for the function.

The rule is written assuming a function with one parameter, but you should generalize it to support multiple parameters; when there are multiple parameters, you should evaluate them in order from left to right. If the caller passes a different number of parameters from what the callee expects, you should throw an error.

The next step, and the most complex one is to allocate a new stack frame for the function call.

The rule specifies this by creating new stack frame with parent pointer equal to  $a$ . The rule also traverses the body of the function and for every global declaration *global x* it should add  $x$  to the set of global variables for the new stack frame. The function *globals* (specified below) performs this traversal. Then, for every assignment of the form  $x = e$  in the body of  $f$ , the rule adds variable  $x$  to the stack frame and initializes it to `None`. The rule *assigns* (specified below) performs this traversal.

After creating the new stack frame, the rule sets all the parameters in the new stack frame to their correct values, and then evaluate the body of the function.

If  $e_1$  does not evaluate to a Function value, then your implementation should return an `IllegalCastException`. If the caller passes too many or too few parameters to the function, then your implementation should return a `RuntimeException`.

## 4 Automatic Casting

Some of the rules above involve automatic casting from other values into strings. Your implementation should define a `toString` function that casts values to strings according to the following rules:

- $\text{toString}(\text{String}(s)) = \text{String}(s)$
- $\text{toString}(\text{Boolean}(b)) = b? \text{"true"} : \text{"false"}$
- $\text{toString}(\text{Integer}(n)) = \text{The string representing the integer } n \text{ in base 10}$
- $\text{toString}(\text{Function}(a, x, s)) = \text{"FUNCTION"}$
- $\text{toString}(\{x_i : v_i\}) = \{\text{"} + x_i + \text{:} + \text{toString}(v_i) + \text{"} \dots \text{"}\}$
- $\text{toString}(\text{None}) = \text{"None"}$

$\text{VARIABLEREAD}$ $\frac{\text{lookup\_read}(a_1, h, x) = a_2 \quad h(a_2) = v}{(\gamma; a_1, h, x) \rightarrow (h, v)}$	$\text{RECORD}$ $\frac{\begin{array}{c} (\gamma, h, e_1) \rightarrow (h_1, v_1) \quad \dots \quad (\gamma, h_{(n-1)}, e_n) \rightarrow (h_n, v_n) \\ \{a_1, \dots, a_{n+1}\} \not\in \text{dom}(h_n) \quad m = \{x_1 : a_1, \dots, x_n : a_n\} \quad h' = h_n[a_1 : v_1, \dots, a_n : v_n, a_{(n+1)} : m] \end{array}}{(\gamma, h, \{x_1 : e_1, \dots, x_n : e_n\}) \rightarrow (h', \text{Record}(a_{(n+1)}))}$
$\text{FIELDREAD}$ $\frac{(\gamma, h, e) \rightarrow (h', \text{Record}(a_1)) \quad h'(a_1) = m \quad x \in \text{dom}(m) \quad a_2 = m(x)}{(\gamma, h, e.x) \rightarrow (h', h'(a_2))}$	$\text{FIELDREADFAIL}$ $\frac{(\gamma, h, e) \rightarrow (h', \text{Record}(a)) \quad h'(a) = m \quad x \notin \text{dom}(m)}{(\gamma, h, e.x) \rightarrow (h', \text{None})}$
$\text{INDEXREAD}$ $\frac{h'(a_1) = m \quad (\gamma, h', e_2) \rightarrow (h'', v_1) \quad x = \text{toString}(v_1) \quad x \in \text{dom}(m) \quad a_2 = m(x)}{(\gamma, h, e_1[e_2]) \rightarrow (h'', h'(a_2))}$	$\text{INDEXREADFAIL}$ $\frac{h'(a) = m \quad (\gamma, h', e_2) \rightarrow (h'', v_1) \quad x = \text{toString}(v_1) \quad x \notin \text{dom}(m)}{(\gamma, h, e_1[e_2]) \rightarrow (h'', \text{None})}$
$\text{FUNCTION}$ $\frac{}{(\gamma; a, h, \text{fun } x \ s) \rightarrow (h, \text{Function}(a, x, s))}$	
$\text{FUNCTIONCALLBASE}$ $\frac{\begin{array}{c} (\gamma, h_1, e_1) \rightarrow (h_2, \text{Function}(a_c, x, s)) \quad (\gamma, h_2, e_2) \rightarrow (h_3, v) \quad a_2 \notin \text{dom}(h_3) \\ a_3 \notin (\text{dom}(h_3) \cup \{a_2\}) \quad \sigma = \{\text{parent} : a_c, \text{global} : (\text{globals}(s), a_g)\}[x_i \in \text{assigns}(s) : a_{\text{None}}][x : a_2] \\ h_4 = h_3[a_2 : v][a_3 : \sigma] \quad (a_g; \gamma; a_3, h_4, s) \rightarrow \eta \end{array}}{(a_g; \gamma; a, h_1, e_1(e_2)) \rightarrow \eta}$	$\text{FUNCTIONCALLRETURN}$ $\frac{(\gamma, h, e_1(e_2)) \rightarrow \text{Return}(h', v)}{(\gamma, h, e_1(e_2)) \rightarrow (h', v)}$
	$\text{FUNCTIONCALLNORETURN}$ $\frac{(\gamma, h, e_1(e_2)) \rightarrow h'}{(\gamma, h, e_1(e_2)) \rightarrow (h', \text{None})}$

Figure 5: Data and Function Expressions

<b>GLOBALLOOKUP</b> $\frac{h(a_1) = \sigma_1 \quad \sigma_1(\text{global}) = (X_g, a_2) \quad x \in X_g \quad h(a_2) = \sigma_2 \quad x \in \text{dom}(\sigma_2) \quad \sigma_2(x) = a_3}{\text{lookup\_read}(a_1, h, x) = a_3}$	<b>LOCALLOOKUP</b> $\frac{h(a_1) = \sigma_1 \quad \sigma_1(\text{global}) = (X_g, a_2) \quad x \notin X_g \quad x \in \text{dom}(\sigma_1) \quad \sigma_1(x) = a_3}{\text{lookup\_read}(a_1, h, x) = a_3}$	<b>SCOPEDLOOKUP</b> $\frac{\sigma_1(\text{global}) = (X_g, a_2) \quad x \notin X_g \quad x \notin \text{dom}(\sigma_1) \quad h(a_1) = \sigma_1 \quad \text{lookup\_read}(\sigma(\text{parent}), x) = a_3}{\text{lookup\_read}(a_1, h, x) = a_3}$
---	---	---

Figure 6: Semantics of the *lookup-read* function

<b>ASSIGN</b> $\overline{\text{assigns}(x = e) = \{x\}}$	<b>HEAPASSIGNMENT</b> $\overline{\text{assigns}(e_1.x = e_2) = \{\}}$	<b>HEAPINDEXASIGNMENT</b> $\overline{\text{assigns}(e_1[e_2] = e_3) = \{\}}$
<b>IF</b>		
$\overline{\text{assigns}(\text{if } e s_1 \text{ else } s_2) = \text{assigns}(s_1) \cup \text{assigns}(s_2)}$		
<b>SEQUENCE</b>		
<b>GLOBAL</b>		
<b>RETURN</b>		
$\overline{\text{assigns}(s_1 ; s_2) = \text{assigns}(s_1) \cup \text{assigns}(s_2)}$		
$\overline{\text{assigns}(\text{global } x) = \{x\}}$		
$\overline{\text{assigns}(\text{return } e) = \{\}}$		

Figure 7: Semantics of the *assigns* function

<b>ASSIGN</b> $\overline{\text{globals}(x = e) = \{\}}$	<b>HEAPASSIGNMENT</b> $\overline{\text{globals}(e_1.x = e_2) = \{\}}$	<b>HEAPINDEXASIGNMENT</b> $\overline{\text{globals}(e_1[e_2] = e_3) = \{\}}$
<b>IF</b>		
$\overline{\text{globals}(\text{if } e s_1 \text{ else } s_2) = \text{globals}(s_1) \cup \text{globals}(s_2)}$		
<b>SEQUENCE</b>		
<b>GLOBAL</b>		
<b>RETURN</b>		
$\overline{\text{globals}(s_1 ; s_2) = \text{globals}(s_1) \cup \text{globals}(s_2)}$		
$\overline{\text{globals}(\text{global } x) = \{x\}}$		
$\overline{\text{globals}(\text{return } e) = \{\}}$		

Figure 8: Semantics of the *globals* function

## 5 Native functions

Your interpreter should support the following native functions:

- **print(s)** Uses the default casting of s to a string and prints it to the console followed by a newline.
- **input()** Reads a line of input from the console and returns it as a string value.
- **intcast(s)** Expects a string and internally uses the c++ function `atoi` to parse the string and return an integer value. If the string does not represent an integer (e.g., the string "hello"), the function should raise an `IllegalCastException`

These three functions should be initially available in the global scope. However, it should be possible for the program to redefine them. For example, the program below updates `print` to add a prefix to every message.

```
print("Hello"); //this should print 'Hello' to the console.  
oldprint = print;  
print = fun(s){  
    oldprint("OUTPUT: " + s);  
};  
print("Hello"); //this should print 'OUTPUT: Hello' to the console.
```