## Homework 9

Due date: Tuesday, November 24, 2015

The primary purpose of this assignment is to explore mutation and imperative updates in programming languages. We explore this language feature together with the related notion of parameter passing modes. Concretely, we extend JAKARTASCRIPT with mutable variables and parameter passing modes. Parameters are always passed by value in JavaScript, so the parameter passing modes in JAKARTASCRIPT are an extension beyond JavaScript to illustrate a language design decision. We will update our type checker and interpreter from Homework 8 and see that mutation forces us to do a rather global refactoring of our interpreter.

Like last time, you will work on this assignment in pairs. However, note that each student needs to submit a write-up and are individually responsible for completing the assignment. You are welcome to talk about these exercises in larger groups. However, we ask that you write up your answers in pairs. Also, be sure to acknowledge those with which you discussed, including your partner and those outside of your pair.

Try to make your code as concise and clear as possible. Challenge yourself to find the most crisp, concise way of expressing the intended computation. This may mean using ways of expressing computation currently unfamiliar to you.

Finally, make sure that your file compiles and runs (using Scala 2.11.7). A program that does not compile will *not* be graded.

Submission instructions. Upload to NYU classes exactly three files named as follows:

- hw09-YourNYULogin.pdf with your answers to the written questions (scanned, clearly legible handwritten write-ups are acceptable).
- hw09-YourNYULogin.scala with your answers to the coding exercises.

Replace YourNYULogin with your NYU login ID (e.g., I would submit hw09-tw47.pdf and so forth). To help with managing the submissions, we ask that you rename your uploaded files in this manner.

Getting started. Download the code pack hw09.zip from the assignment section on the NYU classes page.

## Problem 1 Parameter Passing Modes (12 Points)

This problem is meant as a warm-up exercise for Problem 2 where you will implement a version of JakartaScript that supports different parameter passing modes.

Consider the typed variant of JakartaScript that supports the different parameter passing modes we discussed. For each of the following programs, say whether the program is well-typed. If not, provide a brief explanation of the type error. If yes, compute the value that the program evaluates to and provide a brief explanation why you obtain that specific value.

```
(a)
1 const y = 3;
2 const f = function(var x: number) {
       x = x + 1; return x + y;
     };
f(y) + y
(b)
1 var x = 3;
2 const f = function(name x: number) {
       return x + x;
5 f(x = x + 1) + x
(c)
var y = 3;
2 const f = function(var x: number) {
       return x = x + 1; return x + y
     };
5 f(y * 2) + y
(d)
1 var x = 3;
  const f = function(ref y: number) {
       y = y + 1; return y + x
     };
5 f(x) + x
(e)
var y = 3;
2 const f = function(ref x: number) {
       return x = x + 1; return x + y
     };
5 f(y * 2) + y
```

```
n \in Num
                                                                                           numbers (double)
                      s \in Str
                                                                                                          strings
                      a \in Addr
                                                                                                       addresses
        b \in Bool ::= true \mid false
                                                                                                       Booleans
                      x \in Var
                                                                                                        variables
                      f \in Fld
                                                                                                    field names
          \tau \in \mathit{Typ} ::= \mathbf{bool} \mid \mathbf{number} \mid \mathbf{string} \mid \mathbf{Undefined} \mid
                                                                                                            types
                            (\overline{mode\ x:\tau}) \Rightarrow \tau_0
           v \in Val ::= undefined | n | b | s | a |
                                                                                                           values
                            function p(\overline{mode\ x:\tau})\ t\ e
         e \in Expr ::= x \mid v \mid uop \ e \mid e_1 \ bop \ e_2 \mid e_1 \ ? \ e_2 : e_3 \mid
                                                                                                    expressions
                           console.log(e) \mid e_1(\overline{e}) \mid
                           mut \ x = e_1; e_2
      uop \in Uop ::= - \mid ! \mid \star
                                                                                             unary operators
       bop \in Bop ::= + |-| * | / | === | ! == | < | > |
                                                                                            binary operators
                           <= | >= | & & | | | | =
                   p := x \mid \epsilon
                                                                                              function names
                    t ::= :\tau \mid \epsilon
                                                                                                  return types
      mut \in Mut ::= \mathbf{const} \mid \mathbf{var}
                                                                                                     mutability
mode \in PMode ::= \mathbf{const} \mid \mathbf{var} \mid \mathbf{ref} \mid \mathbf{name}
                                                                                                 passing mode
        M \in Mem = Addr \rightharpoonup Val
                                                                                                      memories
```

Figure 1: Abstract syntax of JakartaScript

## Problem 2 JAKARTASCRIPT Interpreter with State (28 Points)

At this point, we are used to extending our interpreter implementation by updating our type checker typeInfer and our interpreter eval. The syntax with the new extensions highlighted is shown in Figure 1.

Mutation. In this assignment, we add mutable variables declared as follows:

var 
$$x = e_d$$
;  $e_b$ 

and then include an assignment expression:

$$e_1 = e_2$$

that writes the value of  $e_2$  to a location named by expression  $e_1$ .

Parameter Passing Modes. In this assignment, we can annotate function parameters with **const**, **var**, **name**, or **ref** to specify a parameter passing mode. The annotation **var** says that the parameter should be pass-by-value with an allocation for a new mutable parameter variable initialized to the argument value. The **name** and **ref** annotations specify pass-by-name and pass-by-reference, respectively. In Homework 8, all parameters were pass-by-value with an immutable variable. This mode is now captured by a **const** annotation. These "pass-by" terms are defined by their respective EVALCALL rules in Figure 6.

Addresses a and dereference operations  $\star$  a are included in program expressions e because they arise during evaluation. However, there is no way to explicitly write these expressions in the source program. Addresses and dereference expressions are examples of enrichments of program expressions as an intermediate form solely for evaluation.

In Figure 2, we show the updated and new AST nodes. Note that Deref is a Uop and Assign a Bop.

**Type Checking.** The inference rules defining the typing relation are given in Figures 3 and 4. Similar to before, we implement type inference with the function

```
def typeInfer(env: Map[String, (Mut, Typ)], e: Expr): Typ
```

that you need to complete. Note that the type environment maps a variable name to a pair of a mutability (either MConst or MVar) and a type. A template for the Function case for typeInfer is provided that you may use if you wish.

In the implementation of our type checker we now distinguish two different kinds of type errors. Just like in Homework 8, we have regular static type errors for those situations where we detect a type mismatch. The second kind of type error is used to signal the cases where we expect an assignable expression (i.e., a mutable variable) but find a non-assignable expression. To signal this second kind of type error, we will use a Scala exception

```
case class LocTypeError(e: Expr) extends JsException
```

where e is the non-assignable expression. We also provide a helper function locerr to simplify throwing this exception.

For example, in the following program we declare a **const** variable x, which we then try to reassign:

```
const x = 3;
x = 2
```

The function typeInfer should reject this program throwing a LocTypeError with the expression x in the left side of the assignment on line 2 as argument.

**Evaluation.** We also update subst and eval from Homework 8. The required modifications to subst are minimal and have already been provided for you.

A big-step operational semantics is given in Figures 5 and 6. The big-step judgment form is now as follows:

$$\langle M, e \rangle \Downarrow \langle M', v \rangle$$

that says informally, "In input memory M, expression e evaluates to the the output memory M' and value v." The memory M is a map from addresses a to values. The presence of a memory M that gets updated during evaluation is the hallmark of imperative computation.

```
sealed abstract class Expr extends Positional
/** Declarations */
case class Decl(mut: Mut, x: String, ed: Expr, eb: Expr) extends Expr
\texttt{Decl}\,(\mathit{mut}\,,x\,,e_d\,,e_b) \quad \mathit{mut}\,\,x = e_d\,;e_b
sealed abstract class Mut
case object MConst extends Mut
MConst \mathbf{const}
case object MVar extends Mut
MVar var
/** Functions */
type Params = List[(String, (PMode, Typ))]
case class Function(p: Option[String], xs: Params, t: Option[Typ], e: Expr) extends Val
Function (p, \text{ List } ((x, (mode, \tau))), t, e) function p(\overline{mode \ x : \tau}) t e
/* Parameter Passing Modes */
sealed abstract class PMode
case object PConst extends PMode
PConst {f const}
case object PName extends PMode
PName name
case object PVar extends PMode
PVar var
case object PRef extends PMode
PRef \mathbf{ref}
/** Addresses and Mutation */
case object Assign extends Bop
Assign =
case object Deref extends Uop
case class Addr private[ast] (addr: Int) extends Expr
Addr(a) a
/** Types */
sealed abstract class Typ
case class TFunction(ts: List[(PMode, Typ)], tret: Typ) extends Typ
TFunction (List (\overline{(mode, \tau)}), \tau_0) (\overline{mode \tau}) \Rightarrow \tau_0
```

Figure 2: Representing in Scala the abstract syntax of JAKARTASCRIPT. After each case class or case object, we show the correspondence between the representation and the concrete syntax.

Figure 3: Type checking rules for non-imperative primitives of JakartaScript (no changes compared to Homework 8)

$$\frac{\Gamma \vdash e_d : \tau_d \quad \Gamma' = \Gamma[x \mapsto (mut, \tau_d)] \quad \Gamma' \vdash e_b : \tau_b}{\Gamma \vdash mut \ x = e_d; e_b : \tau_b} \quad \text{TypeDecl}$$

$$\frac{x \in \text{dom}(\Gamma) \quad \Gamma(x) = (mut, \tau)}{\Gamma \vdash x : \tau} \quad \text{TypeVar}$$

$$\frac{\Gamma(x) = (\mathbf{var}, \tau) \quad \Gamma \vdash e : \tau}{\Gamma \vdash x = e : \tau} \quad \text{TypeAssignVar}$$

$$\frac{\Gamma \vdash e : (mode_1 \ \tau_1, \dots, mode_n \ \tau_n) \Rightarrow \tau}{\Gamma \vdash e : (mode_i \neq \mathbf{ref} \ \text{or} \ e_i \in Var) \quad \Gamma \vdash e_i : \tau_i} \quad \text{TypeCall}$$

$$\frac{\Gamma' = \Gamma[x_1 \mapsto (mut(mode_1), \tau_1)] \dots [x_n \mapsto (mut(mode_n), \tau_n)]}{\Gamma \vdash e : \tau \quad \tau' = (mode_1 \ \tau_1, \dots, mode_n \ \tau_n) \Rightarrow \tau} \quad \text{TypeFun}$$

$$\frac{\Gamma' = \Gamma[x_1 \mapsto (mut(mode_1), \tau_1)] \dots [x_n \mapsto (mut(mode_n), \tau_n)]}{\Gamma \vdash \text{function} \ (mode_1 \ x_1 : \tau_1, \dots, mode_n \ x_n : \tau_n) e : \tau'} \quad \text{TypeFun}$$

$$\frac{\Gamma' = \Gamma[x_1 \mapsto (mut(mode_1), \tau_1)] \dots [x_n \mapsto (mut(mode_n), \tau_n)]}{\Gamma \vdash \text{function} \ (mode_1 \ x_1 : \tau_1, \dots, mode_n \ x_n : \tau_n) : \tau \ e : \tau'} \quad \text{TypeFunAnn}$$

$$\frac{\Gamma' = \Gamma[x \mapsto \tau'][x_1 \mapsto (mut(mode_1), \tau_1)] \dots [x_n \mapsto (mut(mode_n), \tau_n)]}{\Gamma \vdash \text{function} \ (mode_1 \ x_1 : \tau_1, \dots, mode_n \ x_n : \tau_n) : \tau \ e : \tau'} \quad \text{TypeFunRec}$$

$$\frac{\Gamma \vdash \text{function} \ x \ (mode_1 \ x_1 : \tau_1, \dots, mode_n \ x_n : \tau_n) : \tau \ e : \tau'}{\Gamma \vdash \text{function} \ x \ (mode_1 \ x_1 : \tau_1, \dots, mode_n \ x_n : \tau_n) : \tau \ e : \tau'}} \quad \text{TypeFunRec}$$

$$mut(\text{const}) = mut(\text{name}) = \text{const}$$

$$mut(\text{var}) = mut(\text{ref}) = \text{var}$$

Figure 4: Type checking rules for imperative primitives of JakartaScript

• The eval function now has the following signature

```
def eval(m: Mem, e: Expr): (Mem, Val)
```

corresponding to the updated operational semantics. This function needs to be completed.

Note that the change in the judgment form necessitates updating all rules, even those that do not involve imperative features as in Figure 5. Some rules require allocating fresh addresses. For example, EVALVARDECL specifies allocating a new address a and extending the memory, mapping a to the value  $v_d$  of the defining expression  $e_d$ . The address a is stated to be fresh by the constraint that  $a \notin \text{dom}(M_d)$ . Then  $M_d$  is updated to M' by storing  $v_d$  at address a in  $M_d$ . In the implementation, you call md.alloc(vd) to get a pair (mp, a) consisting of a fresh address a and an updated memory state mp with the location of address a initialized to value vd in md, thus achieving both of the above steps in just one step.

You might notice that in our operational semantics, the memory M only grows and never shrinks during the course of evaluation. Our interpreter only ever allocates memory and never deallocates! This choice is fine in a mathematical model and for this assignment, but a production run-time system must somehow enable collecting garbage—allocated memory locations that are no longer used by the running program. Collecting garbage may be done manually by the programmer (as in C and C++) or automatically by a conservative garbage collector (as in JavaScript, Scala, Java, C#, and Python).

You might also notice that we have a single memory instead of a stack of activation records for local variables and a heap for objects as discussed in Computer Systems Organization. Our interpreter instead simply allocates memory for local variables when they are encountered (e.g., EVALVARDECL). It never deallocates, even though we know that with local variables, those memory locations become inaccessible by the program once the body of a declaration, respectively, function has been evaluation. The key observation is that the traditional stack is not essential for local variables but rather is an optimization for automatic deallocation based on block statements, respectively, function call-and-return.

$$\frac{\langle M, e \rangle \Downarrow \langle M', v \rangle}{\langle M, e \rangle \Downarrow \langle M', h \rangle} \text{ EvalUMinus } \frac{\langle M, e \rangle \Downarrow \langle M', b \rangle}{\langle M, ! e \rangle \Downarrow \langle M', ! b \rangle} \text{ EvalNot}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', \text{true} \rangle}{\langle M, e_1 \& \& e_2 \rangle \Downarrow \langle M'', v_2 \rangle} \text{ EvalAndTrue}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', \text{false} \rangle}{\langle M, e_1 \downarrow \downarrow \langle M', v_2 \rangle} \frac{\langle M', e_2 \rangle \Downarrow \langle M'', v_2 \rangle}{\langle M, e_1 \downarrow \downarrow \downarrow \langle M', v_2 \rangle} \text{ EvalOrFalse}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', \text{false} \rangle}{\langle M, e_1 \downarrow \downarrow \langle M', v_1 \rangle} \frac{\langle M', e_2 \rangle \Downarrow \langle M'', v_2 \rangle}{\langle M, e_1 \downarrow \downarrow \langle M', \text{true} \rangle} \text{ EvalOrTrue}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', \text{false} \rangle}{\langle M, e_1 \downarrow \downarrow \langle M', v_1 \rangle} \frac{\langle M', e_2 \rangle \Downarrow \langle M'', v_2 \rangle}{\langle M, e_1 \downarrow \downarrow \langle M', v_2 \rangle} \text{ EvalSeq}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', v_1 \rangle}{\langle M, e_1 \downarrow \langle M', v_1 \rangle} \frac{\langle M', e_2 \rangle \Downarrow \langle M'', v_2 \rangle}{\langle M, e_1 \downarrow \langle M', v_2 \rangle} \text{ EvalPrint}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', n_1 \rangle}{\langle M, e_1 \downarrow \langle M', n_1 \rangle} \frac{\langle M, e_2 \rangle \Downarrow \langle M'', n_2 \rangle}{\langle M, e_1 + e_2 \downarrow \downarrow \langle M'', n_2 \rangle} \frac{n = n_1 + n_2}{n_1 + n_2} \text{ EvalPlusNum}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', s_1 \rangle}{\langle M, e_1 \downarrow \langle M', s_2 \rangle \Downarrow \langle M'', s_2 \rangle} \frac{\langle M, e_1 \downarrow \downarrow \langle M'', s_2 \rangle}{\langle M, e_1 \downarrow e_2 \downarrow \downarrow \langle M'', n_2 \rangle} \frac{n = n_1 \log n_2 \log p \in \{*, /, -\}}{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', n_2 \rangle} \text{ EvalConstDecl}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', n_1 \rangle}{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', n_2 \rangle} \frac{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', v_2 \rangle}{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', b \rangle} \text{ EvalConstDecl}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', n_1 \rangle}{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', n_2 \rangle} \frac{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', b \rangle}{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', b \rangle} \text{ EvalInequalNum}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', s_1 \rangle}{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', s_2 \rangle} \frac{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', v_2 \rangle}{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', v_2 \rangle} \text{ EvalIfequalStr}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', v_1 \rangle}{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', v_2 \rangle} \frac{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', v_2 \rangle}{\langle M, e_1 \log p e_2 \downarrow \downarrow \langle M'', v_2 \rangle} \text{ EvalIfelse}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', \text{false} \rangle}{\langle M, e_1 \rangle e_1 e_2 \otimes g_1 \Downarrow \langle M'', v_2 \rangle} \text{ EvalIfelse}$$

$$\frac{\langle M, e_1 \rangle \Downarrow \langle M', \text{false} \rangle}{\langle M, e_1 \rangle e_2 \otimes g_1 \Downarrow \langle M'', v_2 \rangle} \text{ EvalIfelse}$$

Figure 5: Big-step operational semantics of non-imperative primitives of JakartaScript. The only change compared to Homework 8 is the threading of the memory.

$$\frac{\langle M,e\rangle \Downarrow \langle M',v\rangle \quad a\in \mathsf{dom}(M')}{\langle M,\star a=e\rangle \Downarrow \langle M'[a\mapsto v],v\rangle} \; \text{EVALASSIGNVAR}$$

$$\frac{a\in \mathsf{dom}(M)}{\langle M,\star a\rangle \Downarrow \langle M,M(a)\rangle} \; \text{EVALDEREF}$$

$$\frac{\langle M,e_d\rangle \Downarrow \langle M_d,v_d\rangle \quad a\notin \mathsf{dom}(M_d)}{\langle M,v_av_av_{di}e_b\rangle \Downarrow \langle M'',v_b\rangle} \; \text{EVALVARDECL}$$

$$\frac{M'=M_d[a\mapsto v_d]}{\langle M,v_av_av_{di}e_b\rangle \Downarrow \langle M'',v_b\rangle} \; \text{EVALVARDECL}$$

$$\frac{\langle M,e_0\rangle \Downarrow \langle M',v_0\rangle \quad v_0 = \mathsf{function} \; v_0 \; (\overline{mode_i}\,x_i\!:\!\overline{\tau_i}):\tau \; e}{\langle M,e_0\rangle \Downarrow \langle M',v_0\rangle \quad v_0 = \mathsf{function} \; (\overline{mode_i}\,x_i\!:\!\overline{\tau_i}):\tau \; e}$$

$$\frac{\langle M,e_0\rangle \Downarrow \langle M',v_0\rangle \quad v_0 = \mathsf{function} \; (const\;x_1\!:\!\tau_i,\overline{mode_i}\,x_i\!:\!\overline{\tau_i}):\tau \; e}{\langle M',e_1\rangle \Downarrow \langle M'',v_0\rangle \quad v_0 = \mathsf{function} \; (\overline{mode_i}\,x_i\!:\!\overline{\tau_i}):\tau \; (e[v_1/x_1]))}$$

$$\frac{\langle M'',v_0'(\overline{e_i})\rangle \Downarrow \langle M''',v_0\rangle}{\langle M,e_0\;(e_1,\overline{e_i})\rangle \Downarrow \langle M''',v_0\rangle} \; \text{EVALCALLCONST}$$

$$\frac{\langle M,e_0\rangle \Downarrow \langle M',v_0\rangle \quad v_0 = \mathsf{function} \; (name\;x_1\!:\!\tau_i,\overline{mode_i}\,x_i\!:\!\tau_i):\tau \; e}{\langle M,e_0\rangle \Downarrow \langle M'',v_0\rangle \quad v_0 = \mathsf{function} \; (name\;x_1\!:\!\tau_i,\overline{mode_i}\,x_i\!:\!\tau_i):\tau \; e}$$

$$\frac{\langle M,e_0\rangle \Downarrow \langle M',v_0\rangle \quad v_0 = \mathsf{function} \; (name\;x_1\!:\!\tau_i,\overline{mode_i}\,x_i\!:\!\tau_i):\tau \; e}{\langle M',e_0\rangle \Leftrightarrow \langle M',v_0\rangle \Leftrightarrow \langle M',v_0\rangle \Leftrightarrow \langle M',v_0\rangle \Leftrightarrow \langle M'',v_0\rangle \Leftrightarrow \langle M'',$$

Figure 6: Big-step operational semantics of imperative primitives of JAKARTASCRIPT.