Homework 9

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.

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.

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 = (let x: Num) => (x = x + 1, x + y);
_{3} f(y) + y
(b)
1 let x = 3;
2 const f = (name x: Num) => x + x;
_3 f(x = x + 1) + x
(c)
1 let y = 3;
2 const f = (let x: Num) => (x = x + 1, x + y);
_3 f(y * 2) + y
(d)
1 let x = 3;
2 const f = (ref y: Num) => (y = y + 1, y + x);
_3 f(x) + x
```

```
(e)

1 let y = 3;
2 const f = (ref x: Num) => (x = x + 1, x + y);
3 f(y * 2) + y
```

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:

let
$$x = e_d$$
; 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, let, name, or ref to specify a parameter passing mode. The annotation let 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 MLet) 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

```
n \in Num
                                                                                         numbers (double)
                      s \in Str
                                                                                                        strings
                      a \in Addr
                                                                                                    addresses
        b \in Bool ::= true \mid false
                                                                                                     Booleans
                      x \in Var
                                                                                                      variables
          	au \in \mathit{Typ} ::= \mathtt{Bool} \mid \mathsf{Num} \mid \mathsf{String} \mid \mathsf{Undefined} \mid
                                                                                                          types
                           (\overline{mode \ \tau}) \Rightarrow \tau_0
          v \in Val ::= undefined \mid n \mid b \mid s \mid a \mid
                                                                                                         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 ::= \texttt{const} \mid \texttt{let}
                                                                                                   mutability
mode \in PMode ::= const \mid let \mid ref \mid name
                                                                                               passing mode
        M \in Mem = Addr \rightharpoonup Val
                                                                                                    memories
```

Figure 1: Abstract syntax of JakartaScript

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:

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

```
/** Mutabilities */
enum Mut:
  case MConst, MLet // <~ const, let</pre>
/** Parameter Passing Modes */
enum PMode:
 case PConst, PName, PLet, PRef // <~ const, name, let, ref</pre>
/** Binary Operators */
enum Bop:
 case Assign // <~ =</pre>
/** Unary Operators */
enum Uop:
  case Deref // <~ *
/** Expressions */
type Params = List[(String, (PMode, Typ))]
enum Expr:
 case Decl(mut: Mut, x: String, ed: Expr, eb: Expr) // <~ mut x = ed; eb</pre>
  case Function(p: Option[String], xs: Params, t: Option[Typ], e: Expr)
   // <~ function p(mode_1 x_1: 	au_1,...,mode_k x_k: 	au_k) t e
  /** Addresses */
  case Addr private[ast] (addr: Int) // <- a</pre>
/** Types */
enum Typ:
  case TFunction(ts: List[(PMode, Typ)], tret: Typ) // <~ (mode_1 \tau_1, \dots, mode_k \tau_k) => \tau_{ret}
```

Figure 2: The abstract syntax of JakartaScript as represented in Scala. After each case, we show the corresponding JavaScript expression represented by that case.

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

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

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

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

$$\frac{\Gamma \vdash e : (mode_1 \ \tau_1, \dots, mode_k \ \tau_k) \Rightarrow \tau}{\Gamma \vdash e (e_1, \dots, e_k) : \tau} \quad \text{TypeCall}$$

$$\frac{\Gamma' = \Gamma[x_1 \mapsto (mut(mode_1), \tau_1)] \dots [x_k \mapsto (mut(mode_k), \tau_k)]}{\Gamma \vdash e : \tau \quad \tau' = (mode_1 \ \tau_1, \dots, mode_k \ \tau_k) \Rightarrow \tau} \quad \text{TypeFunAnn}$$

$$\frac{\Gamma' = \Gamma[x \mapsto \tau'][x_1 \mapsto (mut(mode_1), \tau_1)] \dots [x_k \mapsto (mut(mode_k), \tau_k)]}{\Gamma \vdash (mode_1 \ x_1 : \tau_1, \dots, mode_k \ x_k : \tau_k) \Rightarrow e : \tau'} \quad \text{TypeFunAnn}$$

$$\frac{\Gamma' = \Gamma[x \mapsto \tau'][x_1 \mapsto (mut(mode_1), \tau_1)] \dots [x_k \mapsto (mut(mode_k), \tau_k)]}{\Gamma \vdash e : \tau \quad \tau' = (mode_1 \ \tau_1, \dots, mode_k \ \tau_k) \Rightarrow \tau} \quad \text{TypeFunRec}$$

$$\frac{\Gamma \vdash \text{function} \ x(mode_1 \ x_1 : \tau_1, \dots, mode_k \ x_k : \tau_k) : \tau \ e : \tau'}{\Gamma \vdash \text{function} \ x(mode_1 \ x_1 : \tau_1, \dots, mode_k \ x_k : \tau_k) : \tau \ e : \tau'} \quad \text{TypeFunRec}$$

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

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

Figure 4: Type checking rules for imperative primitives of JakartaScript

that says informally, "In input memory M, expression e evaluates to 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.

• 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, EVALLETDECL 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., EVALLETDECL). 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.

$$\frac{\langle M,e\rangle \Downarrow \langle M,v\rangle}{\langle M,v\rangle} \text{ EVALVAL } \frac{\langle M,e\rangle \Downarrow \langle M',n\rangle}{\langle M,-e\rangle \Downarrow \langle M',-e\rangle} \text{ EVALUMINUS } \frac{\langle M,e\rangle \Downarrow \langle M',b\rangle}{\langle M,!e\rangle \Downarrow \langle M',!b\rangle} \text{ EVALNOT}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,\text{true}\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,v_2\rangle}{\langle M,e_1\&\&e_2\rangle \Downarrow \langle M_2,v_2\rangle} \text{ EVALANDTRUE}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,\text{false}\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,v_2\rangle}{\langle M,e_1\| \vdash e_2\rangle \Downarrow \langle M_2,v_2\rangle} \text{ EVALORFALSE}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,\text{false}\rangle}{\langle M,e_1\&\&e_2\rangle \Downarrow \langle M_1,\text{false}\rangle} \text{ EVALANDFALSE } \frac{\langle M,e_1\rangle \Downarrow \langle M_1,\text{true}\rangle}{\langle M,e_1\| \vdash e_2\rangle \Downarrow \langle M_1,\text{true}\rangle} \text{ EVALORTRUE}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,v_1\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,v_2\rangle}{\langle M,e_1\} \Downarrow \langle M_1,v_1\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,v_2\rangle} \text{ EVALSEQ}$$

$$\frac{\langle M,e\rangle \Downarrow \langle M',v\rangle \quad v \text{ printed}}{\langle M,\text{console.log}(e\rangle) \Downarrow \langle M',\text{undefined}\rangle} \text{ EVALPRINT}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,n_1\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,n_2\rangle \quad n=n_1+n_2}{\langle M,e_1+e_2\rangle \Downarrow \langle M_2,n_2\rangle \quad n=n_1+n_2} \text{ EVALPLUSNUM}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,s_1\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,s_2\rangle \quad s=s_1+s_2}{\langle M,e_1+e_2\rangle \Downarrow \langle M_2,s_2\rangle} \text{ EVALPLUSSTR}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,n_1\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,n_2\rangle \quad n=n_1bop\,n_2 \quad bop\in \{*,*,-\}}{\langle M,e_1bop\,e_2\rangle \Downarrow \langle M_2,n_2\rangle} \text{ EVALCONSTDECL}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,n_1\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,e_2\rangle \quad b=s_1\ bop\,n_2\ bop\in \{*,>=,<,<=\}}{\langle M,e_1\rangle \Downarrow \langle M_1,e_2\rangle \Downarrow \langle M_2,s_2\rangle \quad b=s_1\ bop\,n_2\ bop\in \{*,>=,<,<=\}} \text{ EVALINEQUALNUM}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,s_1\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,s_2\rangle \quad b=s_1\ bop\,n_2\ bop\in \{*,>=,<,<=\}}{\langle M,e_1\rangle \Downarrow \langle M_1,e_2\rangle \Downarrow \langle M_2,s_2\rangle \quad b=s_1\ bop\,n_2\ bop\in \{*,>=,<,<=\}} \text{ EVALINEQUALSTR}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,v_1\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,v_2\rangle}{\langle M,e_1\rangle bop\,e_2\rangle \Downarrow \langle M_2,b\rangle} \text{ EVALIFTHEN}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,v_1\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,v_2\rangle}{\langle M,e_1\rangle bop\,e_2\rangle \Downarrow \langle M_2,v_2\rangle} \text{ EVALIFELSE}$$

$$\frac{\langle M,e_1\rangle \Downarrow \langle M_1,\text{false}\rangle \quad \langle M_1,e_2\rangle \Downarrow \langle M_2,v_2\rangle}{\langle M_2,e_1\rangle e_1,e_2\rangle \Downarrow \langle M_2,v_2\rangle} \text{ EVALIFELSE}$$

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.

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

$$\frac{a \in \mathsf{dom}(M)}{\langle M,*a\rangle \Downarrow \langle M,M(a)\rangle} \text{ EVALDEREFVAR}$$

$$\frac{\langle M,e_d\rangle \Downarrow \langle M_d,v_d\rangle \quad a \notin \mathsf{dom}(M_d)}{\langle M,\mathsf{let} x=v_d;e_b\rangle \Downarrow \langle M'',v_b\rangle} \text{ EVALLETDECL}$$

$$\frac{\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathsf{function} \ x_0 (\overline{mode_i \ x_i:\tau_i}) : \tau \ e}{\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathsf{function} \ x_0 (\overline{mode_i \ x_i:\tau_i}) : \tau \ e} \text{ EVALCALLREC}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathsf{const} \ x_1:\tau_1, \overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e}{\langle M,e_0\rangle \Downarrow \langle M_1,v_1\rangle \quad v_0' = (\overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ EVALCALLCONST}}$$

$$\langle M,e_0(e_1) \Downarrow \langle M_1,v_1\rangle \quad v_0' = (\overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ ev}_0' = (\overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ ev}_0' = (\overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ ev}_0' = (\overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ ev}_0' = (\overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ ev}_0' = (\overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ ev}_0' = (\overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ evalCallName}}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathbf{name} \ x_1:\tau_1, \overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ ev}_0' = (\overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ ev}_0' = (\overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ evalCallName}}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathbf{let} \ x_1:\tau_1, \overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ evalCallName}}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathbf{let} \ x_1:\tau_1, \overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ evalCallLet}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathbf{ref} \ x_1:\tau_1, \overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ evalCallLet}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathbf{ref} \ x_1:\tau_1, \overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ evalCallLet}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathbf{ref} \ x_1:\tau_1, \overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ evalCallLet}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathbf{ref} \ x_1:\tau_1, \overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ evalCallLet}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathbf{ref} \ x_1:\tau_1, \overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ evalCallLet}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathbf{ref} \ x_1:\tau_1, \overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ evalCallLet}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathbf{ref} \ x_1:\tau_1, \overline{mode_i \ x_i:\tau_i} \Rightarrow : \tau \ e} \text{ evalCallLet}$$

$$\langle M,e_0\rangle \Downarrow \langle M_0,v_0\rangle \quad v_0 = \mathbf{ref} \ x_1:\tau_1, \overline$$

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