CSCI 3155: Lab Assignment 4

Fall 2015: Due Saturday, October 24, 2015

The primary purpose of this lab to understand the interplay between type checking and evaluation. Concretely, we will extend JavaScript with immutable objects and extend our small-step interpreter from Lab 3. Unlike all prior language constructs, object expressions do not have an a priori bound on the number of sub-expressions because an object can have any number of fields. To represent objects, we will use collections from the Scala library and thus will need to get used to working with the Scala collection API.

From your team of 8-10 persons in your lab section, find a new partner for this lab assignment (different from your Lab 1 partner). You will work on this assignment closely with your partner. However, note that each student needs to submit and are individually responsible for completing the assignment.

You are welcome to talk about these questions 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.

Recall the evaluation guideline from the course syllabus.

Both your ideas and also the clarity with which they are expressed matter—both in your English prose and your code!

We will consider the following criteria in our grading:

- How well does your submission answer the questions? For example, a common mistake is to give an example when a question asks for an explanation. An example may be useful in your explanation, but it should not take the place of the explanation.

- How clear is your submission? If we cannot understand what you are trying to say, then we cannot give you points for it. Try reading your answer aloud to yourself or a friend; this technique is often a great way to identify holes in your reasoning. For code, not every program that "works" deserves full credit. We must be able to read and understand your intent. Make sure you state any pre-conditions or invariants for your functions (either in comments, as assertions, or as require clauses as appropriate).

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 expression computation currently unfamiliar to you.

Finally, make sure that your file compiles and runs on COG. A program that does not compile will not be graded.
Submission Instructions. Upload to the moodle exactly four files named as follows:

- Lab4-YourIdentiKey.pdf with your answers to the written questions (scanned, clearly legible handwritten write-ups are acceptable)
- Lab4-YourIdentiKey.scala with your answers to the coding exercises
- Lab4Spec-YourIdentiKey.scala with any updates to your unit tests.
- Lab4-YourIdentiKey.jsy with a challenging test case for your JAVASCRIPTY interpreter.

Replace YourIdentiKey with your IdentiKey (e.g., I would submit Lab4-bec.pdf and so forth). Don't use your student identification number. To help with managing the submissions, we ask that you rename your uploaded files in this manner.

Submit your Lab4.scala file to COG for auto-testing. We ask that you submit both to COG and to moodle in case of any issues.

Sign-up for an interview slot for an evaluator. To fairly accommodate everyone, the interview times are strict and will not be rescheduled. Missing an interview slot means missing the interview evaluation component of your lab grade. Please take advantage of your interview time to maximize the feedback that you are able receive. Arrive at your interview ready to show your implementation and your written responses. Implementations that do not compile and run will not be evaluated.

Getting Started. Clone the code from the Github repository with the following command:

```
git clone -b lab4 https://github.com/bechang/pppl-labs.git lab4
```

A suggested way to get familiar with Scala is to do some small lessons with Scala Koans ([http://www.scalakoans.org/](http://www.scalakoans.org/)). Useful ones for Lab 4 are AboutHigherOrderFunctions, AboutLists, AboutPartialFunctions, and AboutInfixPrefixAndPostfixOperators.

1. Feedback. Complete the survey on the linked from the moodle after completing this assignment. Any non-empty answer will receive full credit.

2. Warm-Up: Collections. To implement our interpreter for JAVASCRIPTY with objects, we will need to make use of collections from Scala's library. One of the most fundamental operations that one needs to perform with a collection is to iterate over the elements of the collection. Like many other languages with first-class functions (e.g., Python, ML), the Scala library provides various iteration operations via higher-order functions. Higher-order functions are functions that take functions as parameters. The function parameters are often called callbacks, and for collections, they typically specify what the library client wants to do for each element.

In this question, we practice both writing such higher-order functions in a library and using them as a client.

(a) Implement a function

```scala
def compressRec[A](l: List[A]): List[A]
```
that eliminates consecutive duplicates of list elements. If a list contains repeated elements they should be replaced with a single copy of the element. The order of the elements should not be changed.

Example:
```scala
scala> compressRec(List(1, 2, 2, 3, 3, 3))
res0: List[Int] = List(1, 2, 3)
```

This test has been provided for you in the template.

For this exercise, implement the function by direct recursion (e.g., pattern match on \( l \) and call \( \text{compressRec} \) recursively). Do not call any \( \text{List} \) library methods.

This exercise is from Ninety-Nine Scala Problems:

http://aperiodic.net/phil/scala/s-99/.

Some sample solutions are given there, which you are welcome to view. However, it is strongly encouraged that you first attempt this exercise before looking there. The purpose of the exercise is to get some practice for the later part of this homework. Note that the solutions there do not satisfy the requirements here (as they use library functions). If at some point you feel like you need more practice with collections, the above page is a good resource.

(b) Re-implement the \( \text{compress} \) function from the previous part as \( \text{compressFold} \) using the \( \text{foldRight} \) method from the \( \text{List} \) library. The call to \( \text{foldRight} \) has been provided for you. Do not call \( \text{compressFold} \) recursively or any other \( \text{List} \) library methods.

(c) Implement a higher-order recursive function

```scala
def mapFirst[A](f: A => Option[A])(l: List[A]): List[A]
```

that finds the first element in \( l \) where \( f \) applied to it returns a \( \text{Some}(a) \) for some value \( a \). It should replace that element with \( a \) and leave \( l \) the same everywhere else.

Example:
```scala
scala> mapFirst((i: Int) => if (i < 0) Some(-i) else None)(List(1,2,-3,4,-5))
res0: List[Int] = List(1, 2, 3, 4, -5)
```

(d) Consider again the binary search tree data structure from Lab 1:

```scala
sealed abstract class Tree {
  def insert(n: Int): Tree
}
```

Here, we have implemented the binary search tree \( \text{insert} \) as a method of \( \text{Tree} \). For
this exercise, complete the higher-order method `foldLeft`. This method performs an
in-order traversal of the input tree this calling the callback \( f \) to accumulate a result. Suppose the in-order traversal of the input tree yields the following sequence of data values: \( d_1, d_2, \ldots, d_n \). Then, `foldLeft` yields

\[
f(\cdots(f(f(z, d_1), d_2))\cdots), d_n).
\]

We have provided a test client `sum` that computes the sum of all of the data values in
the tree using your `foldLeft` method.

(e) Implement a function

\[
def strictlyOrdered(t: Tree): Boolean
\]
as a client of your `foldLeft` method that checks that the data values of \( t \) as an in-order
traversal are in strictly ascending order (i.e., \( d_1 < d_2 < \cdots < d_n \)).

Example:

```
scala> strictlyOrdered(treeFromList(List(1,1,2)))
res0: Boolean = false
```

3. **JavaScripty Type Checker**

As we have seen in the prior labs, dealing with conversions and checking for dynamic type
errors complicate the interpreter implementation. Some languages restrict the possible
programs that it will execute to ones that it can guarantee will not result in a dynamic type
error. This restriction of programs is enforced with an analysis phase after parsing known
as *type checking*. Such languages are called strongly, statically-typed. In this lab, we will
implement a strongly, statically-typed version of JavaScript. We will not permit any type
conversions and will guarantee the absence of dynamic type errors.

In this lab, we extend JavaScript with types, multi-parameter functions, and objects/records (see Figure 1). We have a language of types \( \tau \) and annotate function parameters with
types. Functions can now take any number of parameters. We write a sequence of things
using either an overbar or dots (e.g., \( \overline{e} \) or \( e_1, \ldots, e_n \) for a sequence of expressions). An object literal

\[
\{f_1 : e_1, \ldots, f_n : e_n\}
\]
is a comma-separated sequence of field names with initialization expressions surrounded
by braces. Objects in this lab are more like records or C-structs as opposed to JavaScript
objects, as we do not have any form of mutation, dynamic extension, or dynamic dispatch.
The field read expression \( e_1.f \) evaluates \( e_1 \) to an object value and then looks up the field
named \( f \). An object value is a sequence of field names associated with values. The type
language \( \tau \) includes base types for numbers, booleans, strings, and `undefined`, as well as
constructed types for functions \( (\overline{x} : \tau) \Rightarrow \tau' \) and objects \( \{f_1 : \tau_1; \ldots; f_n : \tau_n\} \).

As an aside, we have chosen a syntax that is compatible with the TypeScript language that
adds typing to JavaScript. TypeScript aims to be fully compatible with JavaScript, so it is not
as strictly typed as JavaScript in this lab.
expressions  
\[ e ::= x | n | b | \text{undefined} | uope_1 | e_1 \ bop e_2 | e_1 ? e_2 : e_3 \\
| \text{const } x = e_1; e_2 | \text{console.log}(e_1) \\
| \text{str } | \text{function } p(x : \tau) \ tann e_1 | e_0(x) \\
| \{ f_1 : e_1, \ldots, f_n : e_n \} | e_1.f \]

values  
\[ v ::= n | b | \text{undefined} | \text{str } | \text{function } p(x : \tau) \ tann e_1 \\
| \{ f_1 : v_1, \ldots, f_n : v_n \} \]

unary operators  
\[ uop ::= - | ! \]

binary operators  
\[ bop ::= \ , | + | - | * | / | === | !== | < | <= | > | >= | \&\& | || \]

types  
\[ \tau ::= \text{number } | \text{bool } | \text{string } | \text{Undefined } | (x : \tau) \Rightarrow \tau' | \{ f_1 : \tau_1; \ldots; f_n : \tau_n \} \]

variables  
\[ x \]

numbers (doubles)  
\[ n \]

booleans  
\[ b ::= \text{true } | \text{false} \]

strings  
\[ \text{str} \]

function names  
\[ p ::= x | \varepsilon \]

field names  
\[ f \]

type annotations  
\[ \text{tann} ::= : \tau | \varepsilon \]

type environments  
\[ \Gamma ::= \cdot | \Gamma[x \mapsto \tau] \]

Figure 1: Abstract Syntax of JAVASCRIPTY

statements  
\[ s ::= \text{const } x = e | e | \{ s_1 \} | ; s_1 s_2 \]

expressions  
\[ e ::= \cdots | \text{const } x = e_1; e_2 | (e_1) \\
| \text{function } p(x : \tau) \ tann e_1 | \text{function } p(x : \tau) \ tann \{ s_1 \ \text{return } e_1 \} | (x : \tau) \Rightarrow e \]

Figure 2: Concrete Syntax of JAVASCRIPTY
/* Functions */
case class Function(p: Option[String], params: List[(String, Typ)],
  tann: Option[Typ],
  e1: Expr) extends Expr
  Function(p, (x, τ), tann, e1) function p(x: τ)tann e1
case class Call(e1: Expr, args: List[Expr]) extends Expr
  Call(e1, args) e1(args)

/* Objects */
case class Obj(fields: Map[String, Expr]) extends Expr
  Object(f: e) {f: e}
case class GetField(e1: Expr, f: String) extends Expr
  GetField(e1, f) e1.f

/* Types */
case class TFunction(params: List[(String, Typ)], tret: Typ) extends Typ
  TFunction(x: τ, τ') (x: τ) ⇒ τ'
case class TObj(tfields: Map[String, Typ]) extends Typ
  TObj(f: τ) {f: τ}

Figure 3: Representing in Scala the abstract syntax of JavaScript. After each case class or case object, we show the correspondence between the representation and the concrete syntax.

As before, the concrete syntax accepted by the parser is slightly less flexible than the abstract syntax in order to match the syntactic structure of JavaScript. For function expressions, the body is surrounded by curly braces (i.e., { }) and consists of a statement s1 for const bindings followed by a return with an expression e1. We do, however, support the anonymous function form

\[(x: \tau) \Rightarrow e_1\]

from TypeScript, which is syntactic sugar for

\[\text{function } (x: \tau) \{ \text{return } e_1 \}.\]

In Figure 3, we show the updated and new AST nodes. We update Function and Call for multiple parameters/arguments. Object literals and field read expressions are represented by Object and GetField, respectively.

In this lab, we implement a type checker that is very similar to a big-step interpreter. Instead of computing the value of an expression by recursively computing the value of each sub-expression, we infer the type of an expression, by recursively inferring the type of each sub-expression. An expression is well-typed if we can infer a type for it.

Given its similarity to big-step evaluation, we can formalize a type inference algorithm in a similar way. In Figure 4, we define the judgment form \( \Gamma \vdash e : \tau \) which says informally, “In type environment \( \Gamma \), expression \( e \) has type \( \tau \).” We will implement a function

\[\text{def } \text{typeInfer}(\text{env}: \text{Map}[\text{String}, \text{Typ}], \ e: \text{Expr}): \text{Typ}\]

that corresponds directly to this judgment form. It takes as input a type environment \( \text{env} \) (\( \Gamma \)) and an expression \( e \) (\( e \)) returns a type \( \text{Typ} \) (\( \tau \)). It is informative to compare these rules with the big-step operational semantics from Lab 3.
Figure 4: Typing of JAVASCRIPTY.
The `TypeEquality` is slightly informal in stating

\[ \tau \text{ has no function types.} \]

We intend this statement to say that the structure of \( \tau \) has no function types. The helper function `hasFunctionTyp` is intended to return `true` if a function type appears in the input and `false` if it does not, so this statement can be checked by taking the negation of a call to `hasFunctionTyp`.

To signal a type error, we will use a Scala exception

```scala
case class StaticTypeError(tbad: Typ, esub: Expr, e: Expr) extends Exception
```

where `tbad` is the type that is inferred sub-expression `esub` of input expression `e`. These arguments are used to construct a useful error message. We also provide a helper function `err` to simplify throwing this exception.

We suggest the following step-by-step order to complete `typeInfer`.

1. First, complete the cases for the basic expressions excluding `Function`, `Call`, `Obj`, and `GetField`.
2. Then, work on these remaining cases. These cases use collections, so be sure to complete the previous question before attempting these cases. You can also work on `step` before finishing `typeInfer`.

**Hints:** Helpful library methods here include `map`, `foldLeft`, `zipped`, `foreach`, and `mapValues`. You may want to use `zipped` in the `Call` case to match up formal parameters and actual arguments.

4. **JavaScripty Small-Step Interpreter**

In this question, we update `substitute` and `step` from Lab 3 for multi-parameter functions and objects. Because of type checking, the `step` cases can be greatly simplified. We eliminate the need to perform conversions and should no longer throw `DynamicTypeError`.

We introduce another Scala exception type

```scala
case class StuckError(e: Expr) extends Exception
```

that should be thrown when there is no possible next step. This exception looks a lot like `DynamicTypeError` except that the intent is that it should never be raised. It is intended to signal a coding error in our interpreter rather than an error in the `JAVASCRIPTY` test input.

In particular, if the `JAVASCRIPTY` expression `e` passed into `step` is closed and well-typed (i.e., judgment `\( \vdash e : \tau \)` holds meaning `inferType(e)` does not throw `StaticTypeError`), then `step` should never throw a `StuckError`. This property of `JAVASCRIPTY` is known as **type safety**.

A small-step operational semantics is given in Figure 5. This semantics no longer has conversions compared to Lab 3. It is much simpler because of type checking (e.g., even with the addition of objects, it fits on one page).
As specified, `SEARCHOBJECT` is non-deterministic (why?). As we view objects as an unordered set of fields, it says an object expression takes can take a step by stepping on any of its component fields. To match the reference implementation, you should make the step go on the first non-value as given by the left-to-right iteration of the collection using `Map.find`. We suggest the following step-by-step order to complete `substitute` and `step`.

1. First, import the cases from your Lab 3 `step` function (perhaps excluding `Call`) and simplify them to remove calls to conversions (e.g., `toNumber`) and cases that throw `Dynamic TypeError`. Follow the small-step operational semantics in Figure 5 to see how to simplify your Lab 3 code (e.g., start with `DONEG`).

2. Then, work on the object cases. These are actually simpler than the function cases. **Hint:** Note that field names are different than variable names. Object expressions are not variable binding constructs—what does that mean about `substitute` for them?

3. Then, work on the function cases. **Hints:** You might want to use your `mapFirst` function from the warm-up question. Helpful library methods here include `map`, `foldRight`, `zipped`, and `forall`.
Figure 5: Small-step operational semantics of JavaScript

\[
e \rightarrow e'
\]