

CS143 Compilers 2020 - Written Assignment 3

Due Tuesday, May 26, 2020 at 11:59 PM

This assignment covers semantic analysis, including scoping and type systems. You may discuss this assignment with other students and work on the problems together. However, your write-up should be your own individual work, and you should indicate in your submission who you worked with, if applicable. Assignments can be submitted electronically through Gradescope as a PDF by Tuesday, May 17, 2016 11:59 PM PDT. A L^AT_EX template for writing your solutions is available on the course website.

1. Consider the following Cool programs:

```
1      class A {
2          x: A; -- line 2
3          baz() : A {{x <- new A; x;}}; -- line 3
4          bar(): A {new A}; -- line 4
5          foo(): String {"World!"};
6      };
7      class B inherits A {
8          foo() : String {" "};
9      };
10
11     class C inherits A {
12         foo() : String {"Hello,"};
13     };
14
15     class Main {
16         main(): Object {
17             let io : IO <- new IO, b : B <- new B, c : C <- new C
18             in {{
19                 io.out_string (c.baz().foo());
20                 io.out_string(b.baz().baz().foo());
21                 io.out_string (b.bar().baz().foo());
22             }}
23         };
24     };
```

- (a) What does this code currently print? By changing only the A types on at most lines 2, 3, and 4 to get this program to print "Hello, World!".

```

1  class Main {
2      main (): Object {
3          let io : IO <- new IO, x : Int <- 1 in {
4              {io.out_int (x);
5                  let x : Int <- - 5 in {{
6                      (* x <- YOUR CODE HERE ;*)
7                      io.out_int (x);
8                  }};
9                  if x == 0 then io.out_string ("x") else io.out_int (x)
10                     fi;
11              }}
12  };

```

- (b) Can you replace `(* x <- YOUR CODE HERE ;*)` with at most one line of code containing an assignment to `x` that gets this code to print "10x"? Why or why not?

2. Type derivations are expressed as inductive proofs in the form of trees of logical expressions. For example, the following is the type derivation for $O[Int/y] \vdash y + y : Int$:

$$\frac{\frac{O[Int/y](y) = Int}{O[Int/y], M, C \vdash y : Int} \quad \frac{O[Int/y](y) = Int}{O[Int/y], M, C \vdash y : Int}}{O[Int/y], M, C \vdash y + y : Int}$$

Consider the following Cool program fragment:

```

1      class A {
2          i: Int;
3          j: Int;
4          b: Bool;
5          s: String;
6          o: SELF_TYPE;
7          foo(): SELF_TYPE { o };
8          bar(): Int { 2 * i + j - i / j - 3 * j };
9      };
10     class B inherits A {
11         p: SELF_TYPE;
12         baz(a: Int, b: Int): Bool { a = b };
13         test(c: Object): Object { (* [Placeholder C] *) };
14     };

```

Note that the environments O and M at the start of the method `test(...)` are as follows:

$O = \emptyset[Int/i][Int/j][Bool/b][String/s][SELF_TYPE_B/o][SELF_TYPE_B/p][Object/c][SELF_TYPE_B/self]$

$M = \emptyset[(SELF_TYPE)/(A, foo)][(Int)/(A, bar)][(Int, Int, Bool)/(B, baz)][(Object)/(B, test)]$

For each of the following expressions replacing [Placeholder C], provide the type derivation and final type of the expression, if it is well typed, otherwise explain why it isn't. Assume Cool type rules (you may omit subtyping relationships from the rules when the type is the same, e.g. $Bool \leq Bool$).

(a)

```
1      b <- p.baz(p.bar(), i)
```

(b)

```
1      p <- o.foo()
```

(c)

```
1      b <- baz(i+j, p.bar(i, o.foo()))
```

(d)

```

1      case c of
2          s: Int => s;
3          i: String => j;
4          b: Object => i;
5      esac

```

3. Consider the following Cool program:

```
1      class A {
2          i: Int <- 1;
3          foo(): Int {i};
4      };
5      class B inherits A {
6          baz(): Int {{i <- 2 + i; i;}};
7      };
8      class C inherits B {
9          baz(): Int {{i <- 3 + i; i;}};
10     };
11     class Main {
12         main(): Object {
13             let a : A <- new C, io: IO <- new IO, j : Int in {{
14                 j <- a@B.baz() + a.baz() + a.foo();
15                 io.out_int (j); -- print statement
16             }}
17         };
18     };
```

- (a) This code does not compile. Provide a complete but succinct explanation as to why that is the case. Please note that the error message of `coold` does not count as an explanation, your answer must show that you understand the problem.
- (b) Assume you are given a variant of Cool which is dynamically typed instead of statically typed. Would the behavior of the code above be safe, in the sense of not triggering a runtime type error, in such variant of Cool? Why or why not? If it is safe, what would the print statement output?

4. Consider the following extension to the Cool syntax as given on page 16 of the Cool Manual, which adds arrays to the language:

$$\begin{aligned}
 \text{expr} ::= & \text{new TYPE[expr]} \\
 & | \text{expr[expr]} \\
 & | \text{expr[expr]} < - \text{expr}
 \end{aligned} \tag{1}$$

This adds a new type $T[]$ for every type T in Cool, including the basic classes. Note that the entire hierarchy of array types still has `Object` as its topmost supertype. An array object can be initialized with an expression similar to “`my_array:T[] ← new T[n]`”, where n is an `Int` indicating the size of the array. In the general case, any expression that evaluates to an `Int` can be used in place of n . Thereafter, elements in the array can be accessed as “`my_array[i]`” and modified using a expression like “`my_array[i] ← value`”.

- (a) Provide new typing rules for Cool which handle the typing judgments for: $O, M, C \vdash \text{new } T[e_1], O, M, C \vdash e_1[e_2]$ and $O, M, C \vdash e_1[e_2] < - e_3$. Make sure your rules work with subtyping.
- (b) Consider the following subtyping rule for arrays:

$$\frac{T_1 \leq T_2}{T_1[] \leq T_2[]}$$

This rule means that $T_1[] \leq T_2[]$ whenever it is the case that $T_1 \leq T_2$, for any pair of types T_1 and T_2 .

While plausible on first sight, the rule above is incorrect, in the sense that it doesn't preserve Cool's type safety guarantees. Provide an example of a Cool program (with arrays added) which would type check when adding the above rule to Cool's existing type rules, yet lead to a type error at runtime.

- (c) In the format of the subtyping rule given above, provide the least restrictive rule for the relationship between array types (i.e. under which conditions is it true that $T_1[] \leq T'$ for a certain T' or $T'' \leq T_1[]$ for a certain T'' ?) which preserves the soundness of the type system. The rule you introduce must not allow assignments between non-array types that violate the existing subtyping relations of Cool.
- (d) Add another extension to the language for immutable arrays (denoted by the type $T()$). Analogous to questions 4a and 4c, for this extension, provide: the additional syntax constructs to be added to the listing of page 16 of the Cool manual, the typing rules for these constructs and the least restrictive subtyping relationship involving these tuple types. It is not necessary that this extension interact correctly with mutable arrays as defined above, but feel free to consider that situation.

5. Consider the following assembly language used to program a stack (**r**, **r1**, and **r2** denote arbitrary registers):

```
1      push r: copies the value of r and pushes it onto the stack.
2      top r: copies the value at the top of the stack into r. This
           command does not modify the stack.
3      pop: discards the value at the top of the stack.
4      r1 *= r2: multiplies r1 and r2 and saves the result in r1. r1 may
           be the same as r2.
5      r1 /= r2: divides r1 with r2 and saves the result in r1. r1 may be
           the same as r2. remainders are discarded (e.g., 5 / 2 = 2).
6      r1 += r2: adds r1 and r2 and saves the result in r1. r1 may be the
           same as r2.
7      r1 -= r2: subtracts r2 from r1 and saves the result in r1. r1 may
           be the same as r2.
8      jump r: jumps to the line number in r and resumes execution.
9      print r: prints the value in r to the console.
```

The machine has three registers available to the program: **reg1**, **reg2**, and **reg3**. The stack is permitted to grow to a finite, but very large, size. If an invalid line number is invoked, pop is executed on an empty stack, or the maximum stack size is exceeded, the machine crashes.

- (a) Write code to enumerate the factorial number sequence, beginning with 1! (1, 2, 6, 24, ...), without termination. Assume that the code will be placed at line 100, and will be invoked by setting **reg1**, **reg2**, and **reg3** to 100, 1, and 1 respectively and running '**jump reg1**'. Your code should use the '**print**' opcode to display numbers in the sequence. You may not hardcode constants nor use any other instructions besides the ones given above.
- (b) This 'helper' function is placed at line 1000:

```
1000      push reg1
1001      reg1 -= reg2
1002      reg2 -= reg1
1003      reg3 += reg2
1004      top reg1
1005      pop
1006      jump reg1
```

This 'main' procedure is placed at line 2000:

```
2000      push reg1
2001      push reg3
2002      top reg1
2003      top reg3
2004      pop
2005      pop
2006      jump reg2
2007      print reg3
2008      jump reg2
```

reg1, **reg2**, and **reg3** are set to 0, 1000, and 2000 respectively, and '**jump reg3**' is executed. What output does the program generate? Does it crash? If it does, suggest a one-line change to the helper function that results in a program that does not crash.