μC: A Simple C Programming Language

Compiler 2021 Programming Assignment III
μC Compiler for Java Assembly Code Generation
Due Date: June 24, 2021 at 23:59

Demonstration: to be announced

This assignment involves the code generation part, which is commonly found in modern compilers. Together with the code developed in the first two assignments, in this assignment, your **μC** compiler is expected to generate the Java assembly code (using Jasmin Instructions) with the input **μC** program. The generated code will then be translated to the **Java bytecode** by the **Java assembler**, **Jasmin**. Lastly, the **Java Virtual Machine (JVM)** runs the generated Java bytecode and outputs the expected execution results.

- Environmental Setup
  - Recommended OS: Ubuntu 18.04
  - Install dependencies: `$ sudo apt install flex bison`
  - Java Virtual Machine (JVM): `$ sudo apt install default-jre`
  - Java Assembler (Jasmin) is included in the Compiler hw3 file.

1. Java Assembly Code Generation

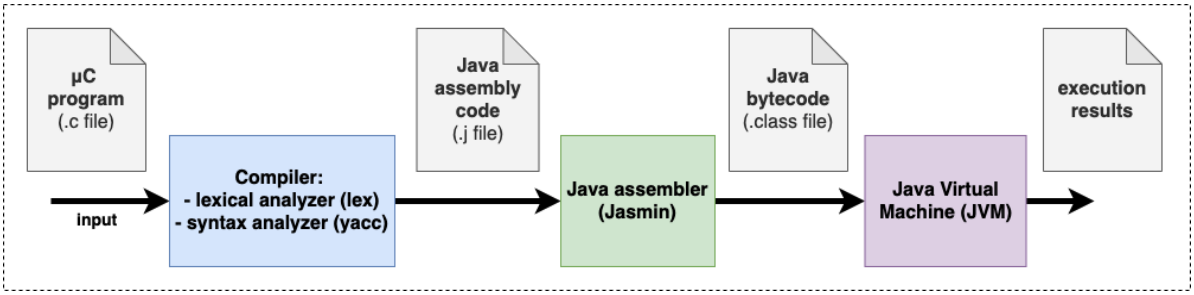

Figure 1. The execution flow for compiling the μC program into Java bytecode for JVM

In this assignment, you have to build a μC compiler. Figure 1 shows the big picture of this assignment and the descriptions for the execution steps are as follows.

- Build your μC compiler by injecting the Java assembly code into your flex/bison code developed in the previous assignments.
- Run the compiler with the given μC program (e.g., test.c file) to generate the corresponding Java assembly code (e.g., test.j file).
- Run the Java assembler, Jasmin, to convert the Java assembly code into the Java bytecode (e.g., test.class file).
- Run the generated Java bytecode (e.g., test.class file) with JVM and display the results.

2. Java Assembly Language (Jasmin Instructions)

In this section, we list the Jasmin instructions that you may use in developing your compiler.

2.1 Literals (Constants)

The table below lists the constants defined in μC language. Also, the Jasmin instructions that we use to load the constants into the Java stack are given. More about the load instructions could be found in the course slides, Intermediate Representation.

Constant in μC	Jasmin Instruction
`94`	`ldc 94`
`8.7`	`ldc 8.7`
`"Hello world"`	`ldc "Hello world"`
`true` / `false`	`iconst_1` / `iconst_0` (or `idc 1` / `ldc 0`)

2.2 Operations

The tables below lists the μC operators and the corresponding assembly code defined in Jasmin (i.e., Jasmin Instruction).

2.2.1 Unary Operators

μC Operator	Jasmin Instruction (`int`)	Jasmin Instruction (`float`)
`+`	- (ignore or a blank)	- (ignore or a blank)
`-`	`ineg`	`fneg`

2.2.2 Binary Operators

μC Operator	Jasmin Instruction (`int`)	Jasmin Instruction (`float`)
`+`	`iadd`	`fadd`
`-`	`isub`	`fsub`
`*`	`imul`	`fmul`
`/`	`idiv`	`fdiv`
`%`	`irem`	-

The following example shows the standard unary and binary arithmetic operations in μC and the corresponding Jasmin instructions.

- μC Code:

```
-5 + 3 * 2;
```

- Jasmin Code (for reference only):

```
ldc 5
ineg
ldc 3
ldc 2
imul
iadd
```

2.2.3 Boolean Operators

μC Operator	Jasmin Instruction
`&&`	`iand`
`||`	`ior`
`!`	`ixor` (`true xor b` equals to `not b`)

- μC Code:

```
// Precedence: ! > && > ||
true || false && !false;
```

- Jasmin Code (for reference only):

```
iconst_1    ; true (1)
iconst_0    ; false (2)
iconst_1    ; load true for "not" operator
iconst_0    ; false (3)
ixor        ; get "not" result (4) from (3)
iand        ; get "and" result (5) from (2),(4)
ior         ; get "or" result from (1),(5)
```

2.2.4 Comparison operators

You need to use subtraction and jump instruction to complete comparison operations. For int, you can use `isub`. For float, there is an instruction `fcmpl` is used to compare two floating-point numbers. Note that the result should be bool type, i.e., 0 or 1. Jump instruction will be mentioned at section 2.6.

- µC Code:

```
1 > 2;
2.0 < 3.1;
```

- Jasmin Code (for reference only):

```
    ldc 1
    ldc 2
    isub
    ifgt L_cmp_0
    iconst_0
    goto L_cmp_1
L_cmp_0:
    iconst_1
L_cmp_1:

    ldc 2.000000
    ldc 3.100000
    fcmpl
    iflt L_cmp_2
    iconst_0
    goto L_cmp_3
L_cmp_2:
    iconst_1
L_cmp_3:
```

2.3 Store/Load Variables

Relative operators: `=`, `+=`, `-=`, `*=`, `/=`, `%=`, `++`, `--`.

2.3.1 Primitive Type

The following example shows how to load the constant at the top of the stack and store the value to the local variable (`x = 9`). In addition, it then loads a constant to the Java stack, loads the content of the local variable, and adds the two values before the results are stored to the local variable (`y = 4 + x`). Furthermore, the example code exhibits how to store a string to the local variable (`z = "Hello"`). The contents of local variables after the execution of the Jasmin code are shown as below.

- µC Code:

```
x = 9;
y = 4 + x;
z = "Hello";
```

- Jasmin Code (for reference only):

```
ldc 9
istore 0        ; store 9 to x

ldc 4
iload 0         ; load x
iadd            ; add 4 and x
istore 1        ; store the result to y

ldc "Hello"
astore 2        ; store a string to z
```

2.3.2 Array Type

The following example shows how to create an variable with array type and store/load the array element. For int array, you need to use `newarray int` to get the reference of an integer array, and `newarray float` for float array. In this assignment, an array can store only integer or floating-point values.

Hint: You may need `swap` instruction to implement array load and store.

- µC Code:

```
int x[3];
int y;
x[0] = 999;
y = x[0] + 4;
```

- Jasmin Code (for reference only):

```
ldc 3           ; array length
newarray int    ; create an array (int32: int, float32: float)
astore 0        ; stroe array variable to local variable 0

ldc 0
istore 1        ; initialize y with 0

aload 0         ; load array
ldc 0           ; index of element
ldc 999         ; value to store to the element
iastore         ; store 999 to the element (x[0])

aload 0         ; load array
ldc 0           ; index of element
iaload          ; load the element (x[0]) to stack
ldc 4
iadd
istore 1        ; store the result to y
```

- Symbol table in this case:

Index	Name	Type	Address	Lineno	Element type
0	x	array	0	1	int
1	y	int	1	2	-

2.4 Print

The following example shows how to print out the constants with the Jasmin code. Note that there is a little bit different for the actual parameters of the println functions invoked by the `invokevirtual` instructions, i.e., int32 (`I`), float32 (`F`), and string (`Ljava/lang/String;`). Note also that you need to treat bool type as string when encountering print statement, and the corresponding code segments are shown as below.

- µC Code:

```
print(30);
print("Hello");
print(true);
```

- Jasmin Code (for reference only):

```
    ldc 30 ; integer
    getstatic java/lang/System/out Ljava/io/PrintStream;
    swap
    invokevirtual java/io/PrintStream/println(I)V

    ldc "Hello" ; string
    getstatic java/lang/System/out Ljava/io/PrintStream;
    swap
    invokevirtual java/io/PrintStream/print(Ljava/lang/String;)V

    iconst_1 ; true
    ifne L_cmp_0
    ldc "false"    ; we should load "false" and "true" as string literal for printing
    goto L_cmp_1
L_cmp_0:
    ldc "true"
L_cmp_1:
    getstatic java/lang/System/out Ljava/io/PrintStream;
    swap
    invokevirtual java/io/PrintStream/print(Ljava/lang/String;)V
```

2.5 Type Conversions (Type Casting)

The following example shows the usage of the casting instructions, `i2f` and `f2i`, where x is int local variable 0, y is float local variable 1.

- µGO Code:

```
x = x + ((int)y)
```

- Jasmin Code (for reference only):

```
iload 0     ; x
fload 1     ; y
f2i         ; convert y to int32
iadd        ; add them
istore 0    ; store to x
```

2.6 Jump Instruction

The following example shows how to use jump instructions (both conditional and non-conditional branches). Jump instruction is used in if statement and for statement.

Jasmin Instruction	Description
`goto <label>`	direct jump
`ifeq <label>`	jump if zero
`ifne <label>`	jump if nonzero
`iflt <label>`	jump if less than zero
`ifle <label>`	jump if less than or equal to zero
`ifgt <label>`	jump if greater than zero
`ifge <label>`	jump if greater than or equal to zero

- µC Code (if statement, x is an int variable):

```
if(x == 10){
    /* do something */
} else {
    /* do the other thing */
}
```

- Jasmin Code (for reference only):

```
    iload 0         ; load x
    ldc 10          ; load integer 10
    isub
    ifeq L_cmp_0    ; jump to L_cmp_0 if x == 0; if not, execute next line
    iconst_0        ; false (if x != 0)
    goto L_cmp_1    ; skip loading true to the stack by jumping to L_cmp_1
L_cmp_0:            ; if x == 0 jump to here
    iconst_1        ; true
L_cmp_1:
    ifeq L_if_false
    ; do something
    goto L_if_exit
L_if_false:
    ; do the other thing
L_if_exit:
```

- µC Code (for statement, x is an int variable):

```
while(x > 0){
    x--;
}
```

- Jasmin Code (for reference only):

```
L_for_begin :
    iload 0             ; x
    ldc 0
    isub
    ifgt L_cmp_0
    iconst_0
    goto L_cmp_1
L_cmp_0 :
    iconst_1
L_cmp_1 :
    ifeq L_for_exit     ; exit when the condition is false
    iload 0             ;---+
    ldc 1               ;   +--- (x--)
```

```
    isub                ;   |
    istore 0            ;---+
    goto L_for_begin    ; goto loop begin
L_for_exit :
```

2.7 Setup Code

A valid Jasmin program should include the code segments for the execution environment setup. Your compiler should be able to generate the setup code, together with the translated Jasmin instructions (as shown in the previous paragraphs). The example code is listed as below.

- Filename: `hw3.j` (generated by your compiler)

```
.source hw3.j
.class public Main
.super java/lang/Object
.method public static main([Ljava/lang/String;)V
.limit stack 100  ; Define your storage size.
.limit locals 100 ; Define your local space number.

    ; ... Your generated Jasmin code for the input µGO program ...

    return
.end method
```

2.8 Workflow Of The Assignment

You are required to build a µC compiler based on the previous two assignments. The execution steps are described as follows.

- Build your compiler by `make` command and you will get an executable named `mycompiler`.
- Run your compiler using the command `$ ./mycompiler < input.c`, which is built by lex and yacc, with the given µC code (`.c` file) to generate the corresponding Java assembly code (`.j` file).
- The Java assembly code can be converted into the Java Bytecode (`.class` file) through the Java assembler, Jasmin, i.e., use `$ java -jar jasmin.jar hw3.j` to generate `Main.class`.
- Run the Java program (`.class` file) with Java Virtual Machine (JVM); the program should generate the execution results required by this assignment, i.e., use `$ java Main.class` to run the executable.

3. What Should Your Compiler Do?

In Assignment 3, the flex/bison file only need to print out the error messages, we score your assignment depending on the JVM execution result, i.e., the output of the command: `$ java Main.class`.

When ERROR occurs during the parsing phase, we expect your compiler to print out ALL error messages, as Assignment 2 did, and DO NOT generate the Java assembly code (.j file).

The total score is 130pt

There 13 test cases which are all included in the Compiler hw3 file.

{"0":"0","1":"30","2":"50","3":"60","4":"70","5":"80","6":"90","7":"100","8":"105","9":"110","10":"115","11":"120","12":"125","13":"130"}

Live Demonstration of Your Assignment 3

To be announced.

4. Submission

- Hand in your homework with Moodle.
- Only allow `.zip` and `.rar` format for compression.
- The directory organization should be:

```
Compiler_StudentID_HW3.zip/
└── Compiler_StudentID_HW3/
    ├── compiler_hw3.l
    ├── compiler_hw3.y
    ├── common.h
    ├── jasmin.jar
    └── Makefile
```

!!! Incorrect format will lose 10pt. !!!

5. References

- The Java Virtual Machine Specification, 2nd Edition (The Java Language Specification, Java SE 12 Edition) https://docs.oracle.com/javase/specs/
- Java bytecode instruction listings (Wikipedia) https://en.wikipedia.org/wiki/Java_bytecode_instruction_listings
- Jasmin homepage http://jasmin.sourceforge.net/