

CSC488 Statements

Jasmeet Sidhu

November 22, 2019

1 Assignment statement:

In an assignment statement, the address of the variable will first be pushed onto the stack, and then the expression will be fully evaluated. Once the expression is fully evaluated, it's final result will be on top of the stack, with the address behind it.

Given the assignment statement

```
1 a := expr
```

the compiler would generate this sequence of instructions:

```
1 ADDR lexical_level offset_of_a
2 (instructions for expression)
3 // expression result now on top of stack
4 STORE
```

2 If statement:

In if statements, the condition will be evaluated first, and its result will be on top of the stack. Then the address to the end of the if body will be pushed onto the stack, and finally a BF instruction will be used to skip the if body if the condition is false.

Given the if statement

```
1 if expression then statement
```

the compiler would generate this sequence of instructions:

```
1 (instructions to evaluate condition)
2 // condition result now on top of stack
3 PUSH address_to_end_of_if_body
4 BF
5 (instructions for if statement body)
6
7 // end of if body
```

If there is an else statement, then the compiler will need to insert a branch in the if body to skip the else body. Also, if the condition is not satisfied, then the compiler must branch into the else body.

Given the if and else statements

```
1 if expression then
2     statement
3 else
4     statement
```

the compiler should generate this sequence of instructions

```
1 (instructions to evaluate condition)
2 // condition result now on top of stack
3 PUSH address_of_else_body
4 BF
5 (if_body)
6 PUSH address_to_end_of_else_body // skip the else body
7 BR
8
9 // else body start
10 (else_body)
11
12 // end of else body
```

3 Loop statements

Similarly to the if statements, the while loop will evaluate its condition first, and so the expression result will be on top of the stack. Then the program can simply branch to skip the loop body if the condition is not met. Additionally, a branch instruction will be placed at the end of the loop body to jump to the beginning of the loop.

Given the while loop

```
1 while expression do
2     statement
```

the compiler should generate this sequence of instructions:

```
1     // loop start
2     (instructions to evaluate the condition)
3     // condition result now on top of stack
4     PUSH address_to_end_of_loop_body
5     BF
6     (loop body)
7     PUSH address_to_loop_start
8     BR
9
10    // end of loop body
```

A similar strategy is used for the repeat statement. The only difference is that the condition is evaluated after the loop body. Given the repeat loop

```
1 repeat
2     statement
3 until expression
```

the compiler should generate this sequence of instructions:

```
1     // loop start
2     (loop body)
3     (instructions to evaluate condition)
4     // condition result now on top of stack
5     PUSH address_to_end_of_loop // jump out of loop if condition not satisfied
6     BF
7     PUSH address_to_loop_start // jump back to start of loop
8     BR
9
10    // end of loop
```

4 Return statements

Since the activation record is cleaned by the caller, a procedure return statement simply needs to read the activation record's return address and then branch to it.

Function return statements are very similar to procedural return statements, except they need to write their return value into the reserved return value section of the activation record. After the return value is written, the function can branch to the return address from the activation record.

5 Exit statements

For a simple exit statement of the form

```
1      exit
```

the compiler should simply branch out of the current loop:

```
1      PUSH address_to_end_of_current_loop
2      BR
```

When given the following conditional multi-loop exit statement

```
1      exit when expression
```

the compiler must negate the result of the condition and then branch if false:

```
1      (instructions to evaluate the condition)
2      // condition result now on top of stack
3      PUSH MACHINE_FALSE
4      EQ
5      PUSH address_to_end_of_current_loop
6      BF
```

The multi-loop exit instructions are very similar. Given the following multi-loop exit statement

```
1      exit N
```

the compiler should generate these instructions

```
1      PUSH address_to_end_of_Nth_loop
2      BR
```

Similarly, when given the following conditional multi-loop exit statement

```
1      exit N when expression
```

the compiler must generate:

```
1      (instructions to evaluate the condition)
2      // condition result now on top of stack
3      PUSH MACHINE_FALSE
4      EQ
5      PUSH address_to_end_of_Nth_loop
6      BF
```

6 Minor scopes

At the beginning of the a minor scope, stack space is allocated for all the local variables. Similarly, the stack space is free'd at the end of the minor scope.

Hence, given a minor scope

```
1      {
2          (declarations)
3          (statements)
4      }
```

the compiler would output

```
1      PUSH 0
2      DUPN (local_stack_size - 1)
3
4      (instructions for statements)
5
6      POPN local_stack_size
```

If there are no local variables, meaning the local stack size is zero, there would be no stack space allocation/deallocation. Hence the generated code would look like as if theres no minor scope:

```
1      (instructions for statements)
```

7 I/O statements

Even though the virtual machine supports reading characters, the language only allows integers to be read. Hence read instructions only support integers.

Given a read integer(s) statement

```
1      read a, b, c, ...
```

the compiler would generate the following instructions

```
1      ADDR lexical_level offset_of_a
2      READI
3      STORE
4
5      ADDR lexical_level offset_of_b
6      READI
7      STORE
8
9      ADDR lexical_level offset_of_c
10     READI
11     STORE
12
13     ...
```

Write instructions support both integers and strings.

Given a string printing statement

```
1      write str
```

the compiler generates the following instructions

```
1      PUSH chr1 // push first character of string
2      PRINTC
3
4      PUSH chr2 // push second character of string
5      PRINT C
6
7      ...
```

Similarly, when given a "write newline" statement, the compiler generates

```
1      PUSH '\n' // push newline character
2      PRINTC
3      ...
```

Integer/expression printing statements are also identical.

Given an expression printing statement

```
1      write expr
```

the compiler generates

```
1      (code to evaluate the expression)
2      // result is on top of stack
3      PRINTI
```

When there are multiple write statements, the compiler simply performs each write statement sequentially like read.