

[TA only: Remember to Announce]

- L1 Tests due Tuesday 9/14. Submissions for coding assignments are through Notolab.
- L1 Checkpoint due Tuesday 9/14. L1 Compiler due Tuesday 9/21.
- Written 1 due Tuesday 9/21. Submissions are through Gradescope.
- Give a brief description of stages of the compiler required for L1 checkpoint and compiler.

Lab 1

Designing a compiler for each of the labs is the heart of this course. The handouts for Lab 1 checkpoint and Lab 1 compiler are both linked on the course website at <https://www.cs.cmu.edu/~411/>, under assignments. Lab submissions are via Notolab at <https://notolab.cs.cmu.edu/>. In very short summary,

- Lab 1 checkpoint asks you to implement register allocation based on graph coloring algorithm given in class.
- Lab 1 asks you to implement the compiler for the language L1.

This semester, we are providing a virtual machine to execute abstract assembly. If you output programs in this form after each phase of compilation, it will be a valuable debugging tool! The code for the VM is available at <https://github.com/15-411/vm> along with examples for L1.

Read the handouts carefully and good luck have fun!

Tracing Through the Backend

In this recitation, we're going to discuss an example of the processing done by the compiler backend. Since you won't have to touch the frontend for Lab 1 (unless you opt to not use any starter code), we'll leave it for a future week. Here's the code and AST we'll use for the example:

```

1 int main() {
2   int x = 42;
3   int z;
4   if (x % 2 == 0) {
5     x++;
6     z = 1;
7   } else {
8     z = -1;
9   }
10  int y = 1;
11  while (y <= x - 1) {
12    y = x + y;
13  }
14  return z * y;
15 }

```

```

1 declare(x, seq(
2   assign(x, const(42)),
3   declare(z, seq(
4     if(compare(mod(x, const(2)), const
5       (0), EQ), seq(
6         incr(x),
7         assign(z, const(1))
8       ),
9     assign(z, neg(const(1)))
10    ),
11  declare(y, seq(
12    assign(y, const(1)), seq(
13      while(
14        compare(y, minus(x, 1),
15          LEQ),
16        assign(y, add(x, y))

```

```

15         ),          18         ))
16         return(times(z,y))    19     ))
17     )                20 ))

```

Code Generation

As discussed in lecture yesterday, we use the "maximal munch" algorithm to generate abstract 3-address assembly from IR. For each line in the IR, we recursively pattern-match as deep as possible into each sub-expression and generate lines of assembly at each step.

Recall that one formulation of the rules we used in lecture was the following:

e	$\text{codegen}(d, e)$	up
c	\cdot	c
x	\cdot	x
$e_1 \oplus e_2$	$t \leftarrow \text{codegen}(e_1) \oplus \text{codegen}(e_2)$	t

and for statements:

s	$\text{codegen}(s)$
$x = e$	$x \leftarrow \text{codegen}(e)$
return e	$\%rax \leftarrow \text{codegen}(e), \text{ret}$

Checkpoint 0

Translate the above AST into abstract 3-address assembly. You do not need to turn it into SSA form here. NOTE: answers may differ depending on the exact rules you're using in your head to generate instructions.

```

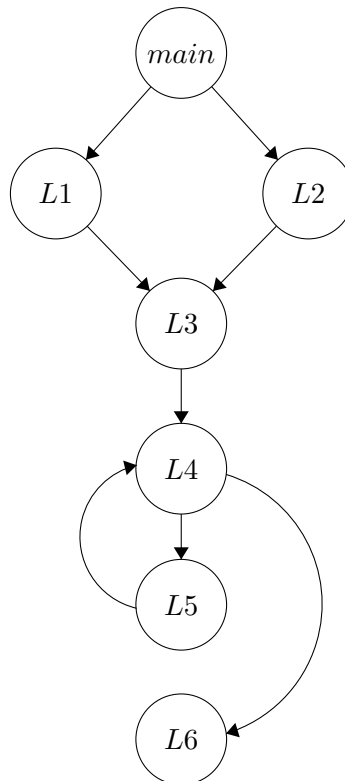
1 main:
2  x ← 42
3  t1 ← 2
4  t2 ← x % t1
5  if (t2 == 0) then goto L1 else goto L2
6 L1:
7  x ← x + 1
8  z ← 1
9  goto L3
10 L2:
11 t3 ← 1
12 z ← t3 * -1
13 goto L3
14 L3:
15 y ← 1
16 L4:
17 t4 ← x - 1
18 if (y <= t4) then goto L5 else goto L6
19 L5:
20 t5 ← x + y
21 y ← t5
22 goto L4

```

```
23 L6:  
24 %eax ← z * y  
25 return
```

Checkpoint 1

Draw the control flow graph (CFG) based on the 3-address assembly. Think about where you need to put Φ functions if you turn it into SSA form.



```
1 main:  
2 x0 ← 42  
3 t1 ← 2  
4 t2 ← x0 % t1  
5 if (t2 == 0) then goto L1 else goto L2  
6 L1:  
7 x1 ← x0 + 1  
8 z0 ← 1  
9 goto L3  
10 L2:  
11 t3 ← 1  
12 z1 ← t3 * -1  
13 goto L3  
14 L3:  
15 x2 ← phi(x1, x0)  
16 z2 ← phi(z0, z1)  
17 y0 ← 1
```

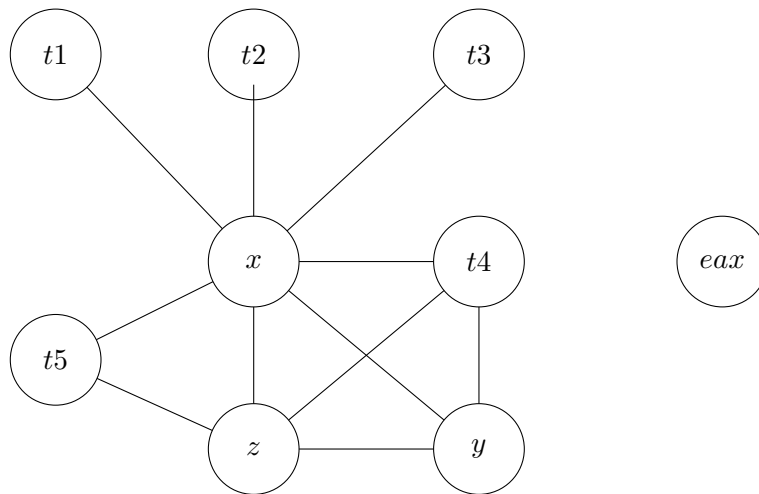
```

18 L4:
19  y1 ← phi(y0, y2)
20  t4 ← x2 - 1
21  if (y1 ≤ t4) then goto L5 else goto L6
22 L5:
23  t5 ← x2 + y1
24  y2 ← t5
25  goto L4
26 L6:
27  %eax ← z2 * y1
28  return

```

Maximum Cardinality Search

From the last recitation, you already know how to perform liveness analysis and construct interference graph. For this program, the graph is given here (as exercise, you could construct the graph yourself):



In order to color the interference graph using the greedy algorithm, we need to decide on an order in which to process the vertices. We do this using the Maximum Cardinality Search algorithm. We first assign a weight of 0 to each vertex. Then, at each step, we:

- (a) Choose a vertex with maximal weight from the working set
- (b) Add it to our ordering and remove it from the working set
- (c) Increment the weights of all of its neighbors

This algorithm produces an ordering which is optimal for chordal graphs.

Checkpoint 2

Use Maximum Cardinality Search to generate an ordering of the vertices in the example above. Break ties by choosing the vertex that is lexicographically first.

```
%eax, t1, x, t2, t3, t4, y, z, t5
```

Greedy Graph Coloring

Once we have an ordering, we can assign registers to each of the temps in our program. Ignoring pre-colored vertices, such as `%eax`, we can color the temps by assigning the lowest register that is not assigned to any of the vertex's neighbors.

Checkpoint 3

Perform Greedy Graph Coloring on the interference graph from above to assign registers `%r1`, `%r2`, ... to the temps in the program. Then rewrite the abstract assembly using the new registers.

Mapping:

```
t1 => %r1
x  => %r2
t2 => %r1
t3 => %r1
t4 => %r1
y  => %r3
z  => %r4
t5 => %r1
```

Program:

```
1 main:
2  %r2 <- 42
3  %r1 <- 2
4  %r1 <- %r2 % %r1
5  if (%r1 == 0) then goto L1 else goto
    L2
6 L1:
7  %r2 <- %r2 + 1
8  %r4 <- 1
9  goto L3
10 L2:
11 %r1 <- 1
12 %r4 <- %r1 * -1
13 goto L3
14 L3:
15 %r3 <- 1
16 L4:
17 %r1 <- %r2 - 1
18 if (%r3 <= %r1) then goto L5 else
    goto L6
19 L5:
20 %r1 <- %r2 + %r3
21 %r3 <- %r1
22 goto L4
23 L6:
24 %eax <- %r4 * %r3
25 return
```

Lab 1 Tip: Spilling Temps

We can't fit all of our data in registers, so we spill into memory. But we need at least one operand in a register for most arithmetic operations. This is getting into the software engineering part of the course, but we will outline one strategy that you can use.

You will need to reserve a register, typically `%r11d`. Perform register allocation, then scan through your instructions looking for memory-memory operations. You then insert a `mov` from the destination to `%r11d`, perform the operation, then move `%r11d` back to memory.

In a functional language, you can implement this in a pass similar to code generation, where you case

on instruction type and produce either a list with the input instruction, or a list with the moves into and out of %r11d.