Sarah Huang
COSC 461: Intro to Compilers
November 22, 2022

Csem Project Report

1. The problem I attempted to solve is to read a snippet of C code and convert into LLVM intermediate representation. It should support variables, loops, conditional statements, functions, goto, labels, etc. A lot of the infrastructure was already set up, so all I needed to do was implement the functions in sem.cpp.

2. A good tip to figure out where to start was building the example inputs with csem_ll. The csem_ll target will stop compilation after running csem if it does not generate the correct LLVM intermediate representation yet. All of the incompleted functions that are called will print out that they have not been implemented.

```
make csem_ll INPUT=input2
...
sem: set not implemented
sem: set not implemented
sem: set not implemented
sem: set not implemented
sem: genstring not implemented
...(and so on)
```

    The way I tackled all of these functions was to work incrementally with a bit of trial and error. I would work on, for example, the set function first and build it, so I see if I was generating the right LLVM intermediate representation compared to the ref_csem. Once it wrote it correctly, I would move on to genstring and continue to work incrementally until the ref_input.ll and csem_input.ll files are the same. I would do the same for all of the input files.

3. To check the program, I built all the example inputs with csem and ref_csem. I then compared the ref_input.ll and csem_input.ll files to see if they were the same.

```
make csem_build INPUT=input2
make ref_build INPUT=input2
Vimdiff ref_input2.ll csem_input2.ll
```

4. The sem.cpp is very long with lots of functions and comments, so it was difficult to navigate. I struggled with deciphering sem.cpp because I did not write everything in

sem.cpp, and the sheer volume did not make it easy. Additionally, I am not familiar with the LLVM IRBuilder and the online descriptions of what each method does were challenging for me to understand. Lastly, and this was definitely on me, I did not give myself enough time to complete this assignment. We had many weeks to do it, and I pushed it off to start it during the last possible weekend. The same weekend I had to go to the NSLDC conference in Newark, NJ. Very smart idea, Sarah. Also, seg fault errors are the bane of my existence.