

## Project 4

### Small C compiler for LLVM IR

The goal of our project this semester is to develop a prototype C compiler. This is done by compiling C programs into embedded platforms such as 68K, MIPS, ARM, or pseudo assemblies. In this class we will use **LLVM IR** (<https://llvm.org/docs/LangRef.html>) as the target codes. **The project is divided into several parts including language definition, lexical analyzer, C-grammar, symbol table handlings, parser, and code generation.** In the final assignment, you will need to choose proper language features which you want to develop in this course, and complete the whole compiler (including the code generator).

In your hand-in report, you need to have the followings:

- **To define the subset of the language which you want to choose from C.**
  - Give a set of testing programs and the generated LLVM-IR files, which can illustrate the features of your testing programs. (at least 3 test programs)
  - Use the “**ANTLR**” to help you develop your compiler.
  - You can use **C**, **C++**, or **Java** to write your compiler. (Java is recommended)
  - Your C subset should at least contain the following features:
    - Integer data types: `int`
    - Statements for arithmetic computation. (e.g., `a=b+2*(100-1);`)
    - Comparison expression. (e.g., `a > b`) , comparison operation: `>` 、 `>=` 、 `<` 、 `<=` 、 `==` 、 `!=`
    - if-then / if-then-else program constructs.
    - `printf()` function with one/two parameters. (support types: `%d`)
- For example:
- `printf("Number is %d\n", var);`
  - `printf("Hello\n");`

可擴充的其他功能:

- Nested if construct
- For-loop construct/ while-loop construct
- Switch-case construct
- Loop construct + if construct
- Floating point data type
- Structure data type
- Logical operation
- Function call
- ...

Please turn in the following files to eCourse2:

- A file describes your language which is a subset of C language. (MS-WORD file)
- The source codes:
  - Your compiler
    - (a) ANTLR grammar file, myCompiler.g.
    - (b) A program to call your parser, myCompiler\_test.java.
  - Testing programs (.c) and the corresponding LLVM IR codes (.ll) generated by your compiler. (at least 3 programs)
- A readme file (pure text file) describes the features of your compiler and explains how to compile and execute your compiler.
- Makefile
- Any other required materials

請將上述的檔案打包壓縮存成 **tar.xz** 之格式，檔名為 **prj4\_<學號>.tar.xz** 檔（例如：**prj4\_401234567.tar.xz**，上傳到 eCourse2。

**Due Date: June 16 (Friday), 24:00, 2023 (無法補交)**