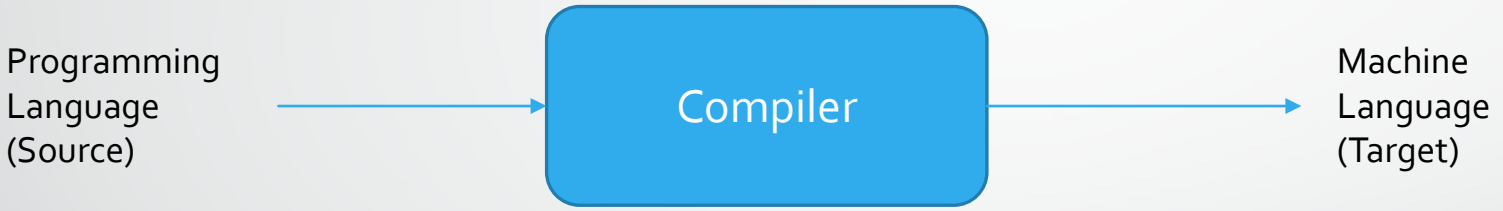
How the compiler works

By: Jorge Luis Rodriguez

What is a Compiler?

- It is a program that translates human readable source code into computer executable machine code.


Programming
Language
(Source)
Compiler
Machine
Language
(Target)

What Do Compilers Do?

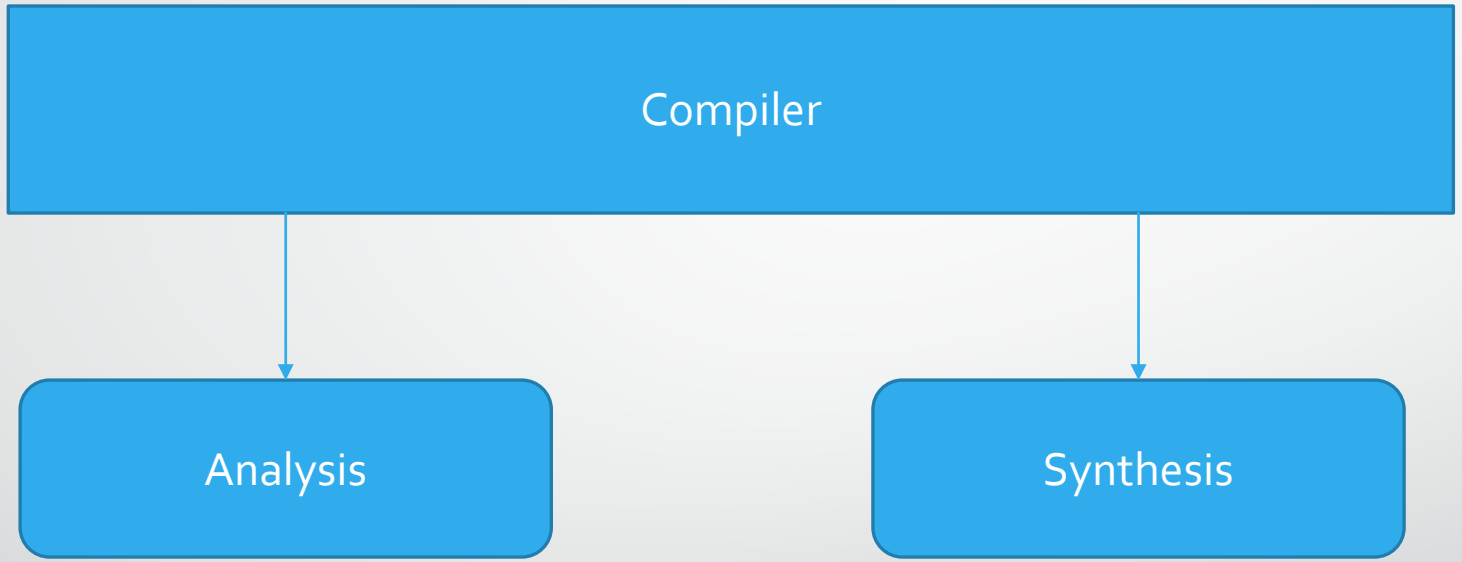
- They may generate three types of code:
  - Pure Machine Code
    - Machine instruction set without assuming the existence of any operating system or library.
    - Mostly being OS or embedded applications.
  - Augmented Machine Code
    - Code with OS routines and runtime support routines.
    - More often.
  - Virtual Machine Code
    - Virtual instructions, can be run on any architecture with a virtual machine interpreter or just-in-time compiler.
    - Example Java.

What Do Compilers Do

- What make compiler differ from one another is in the format of the target machine code they generate:
  - Assembly or another source format
  - Re-locatable binary
    - Relative address
    - A linkage step is required
  - Absolute binary
    - Absolute address
    - Can be executed directly.

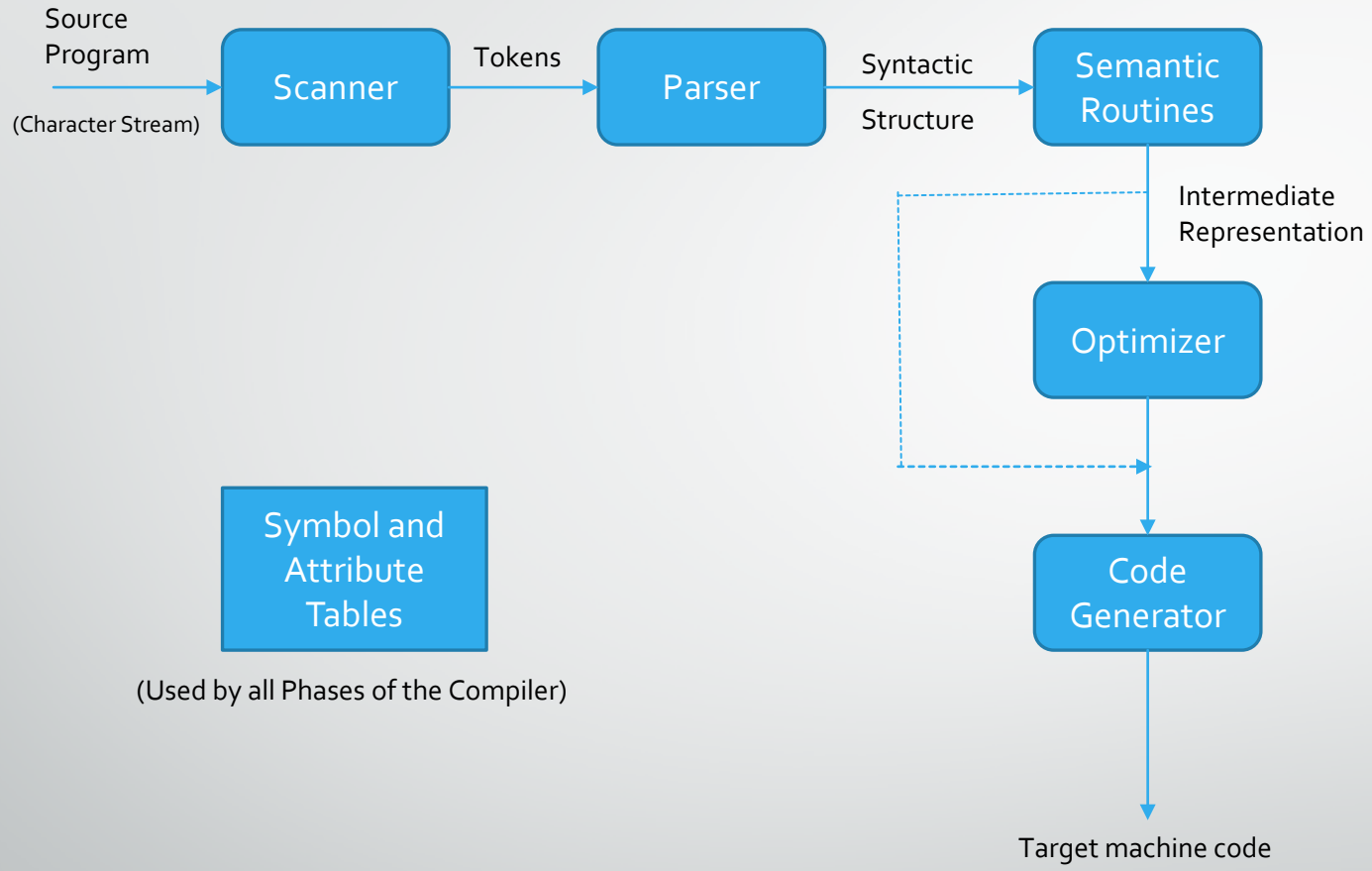
The Structure of a Compiler

- Any compiler must perform two major tasks:
  - Analysis of the source program
  - Synthesis of a machine-language program

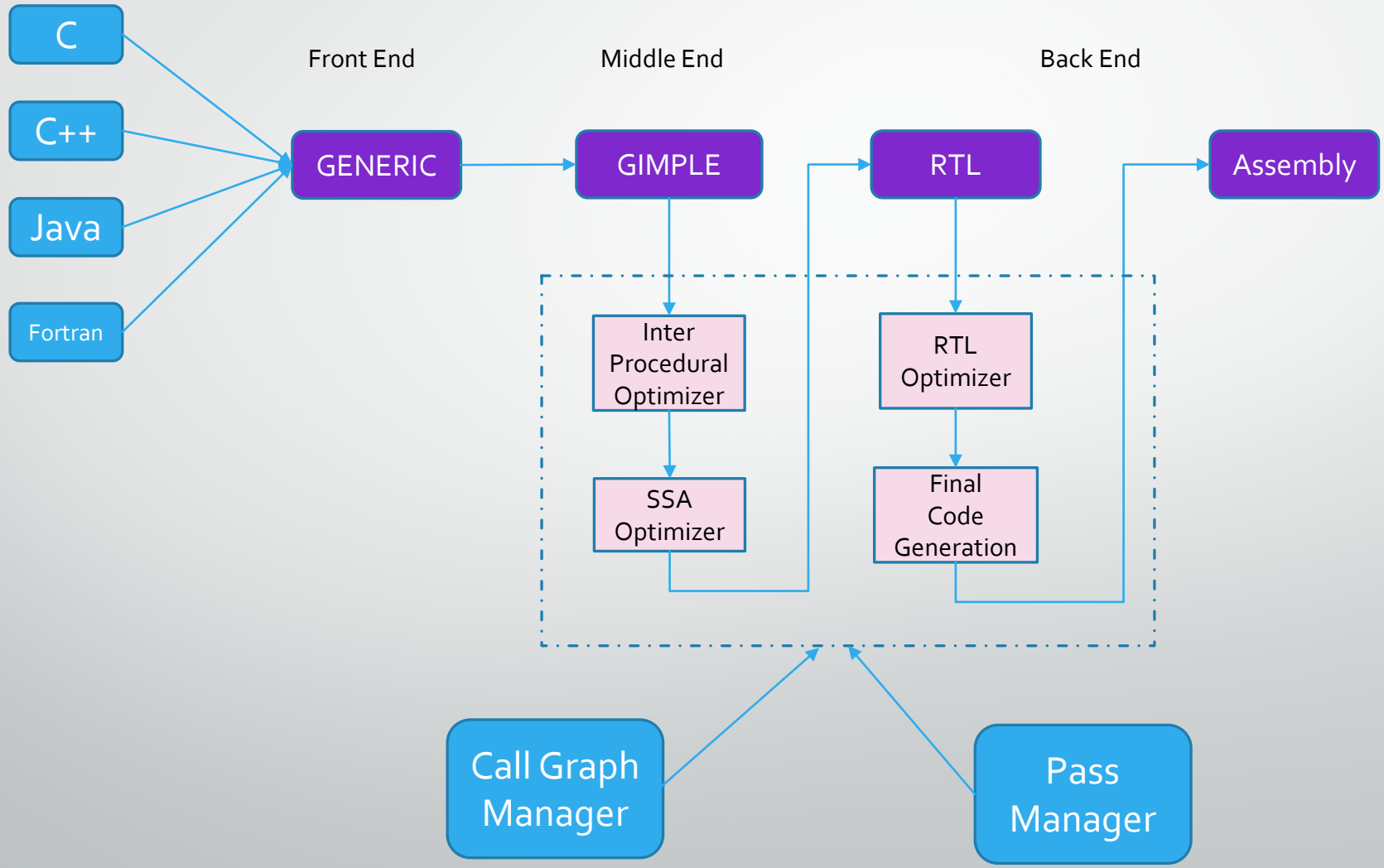

Compiler
Analysis
Synthesis

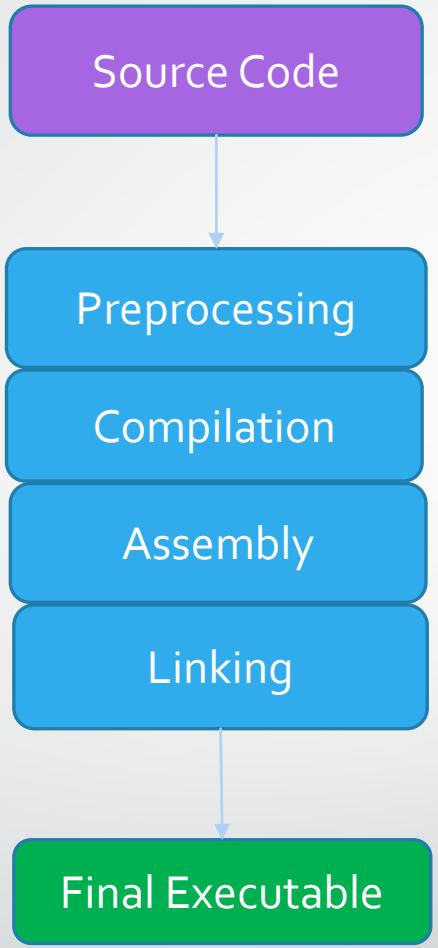
The Structure of a Compiler

- Lexical Analysis
- Parsing
- Semantic Analysis
- Optimization
- Code Generation


Source Program
(Character Stream)
Scanner
Tokens
Parser
Syntactic Structure
Semantic Routines
Intermediate Representation
Optimizer
Code Generator
Target machine code
Symbol and Attribute Tables
(Used by all Phases of the Compiler)

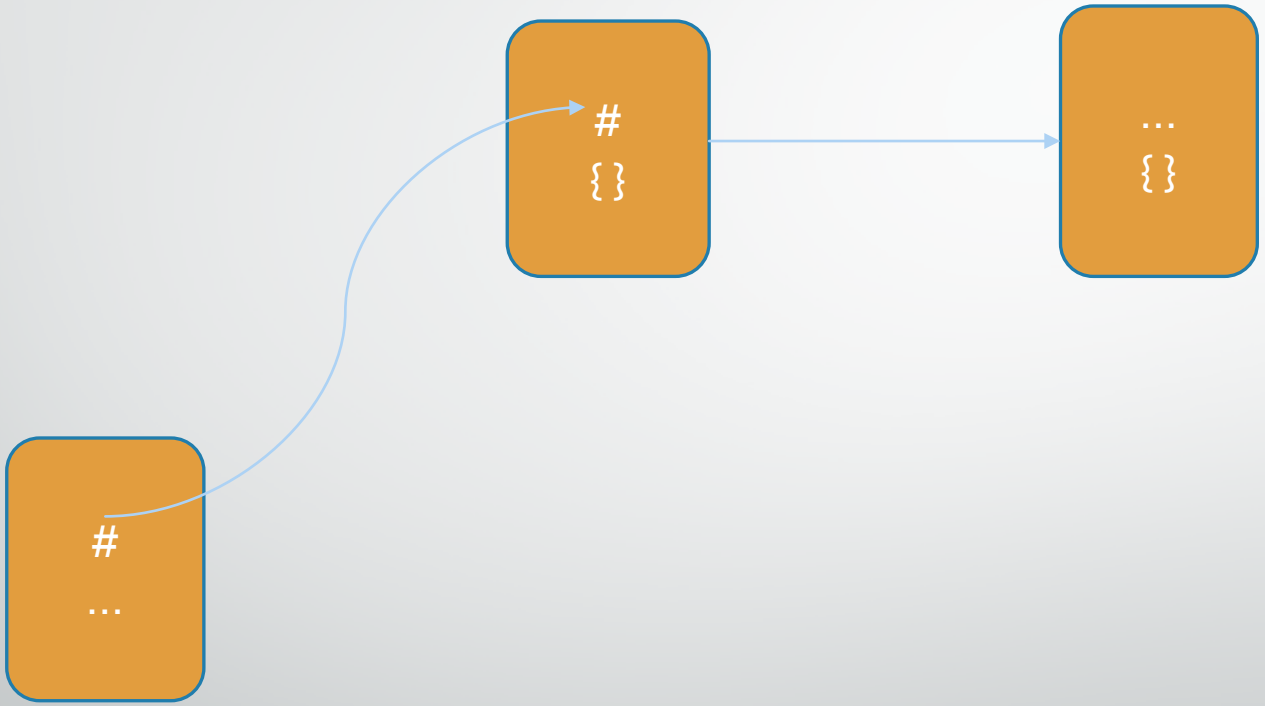
GCC Compiler Overview




Source Code
Preprocessing
Compilation
Assembly
Linking
Final Executable

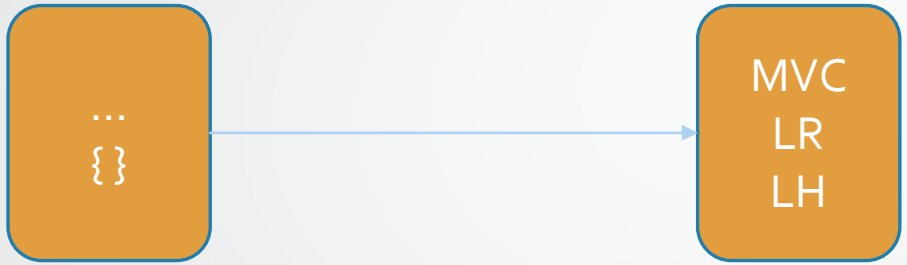
Preprocessing

Expand Macros



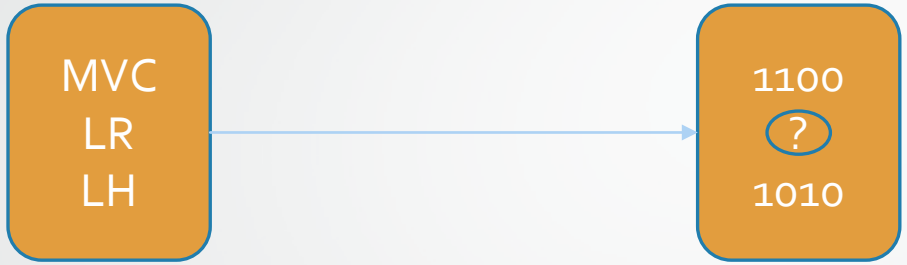
Compilation

From Source Code to Assembly language



Assembly

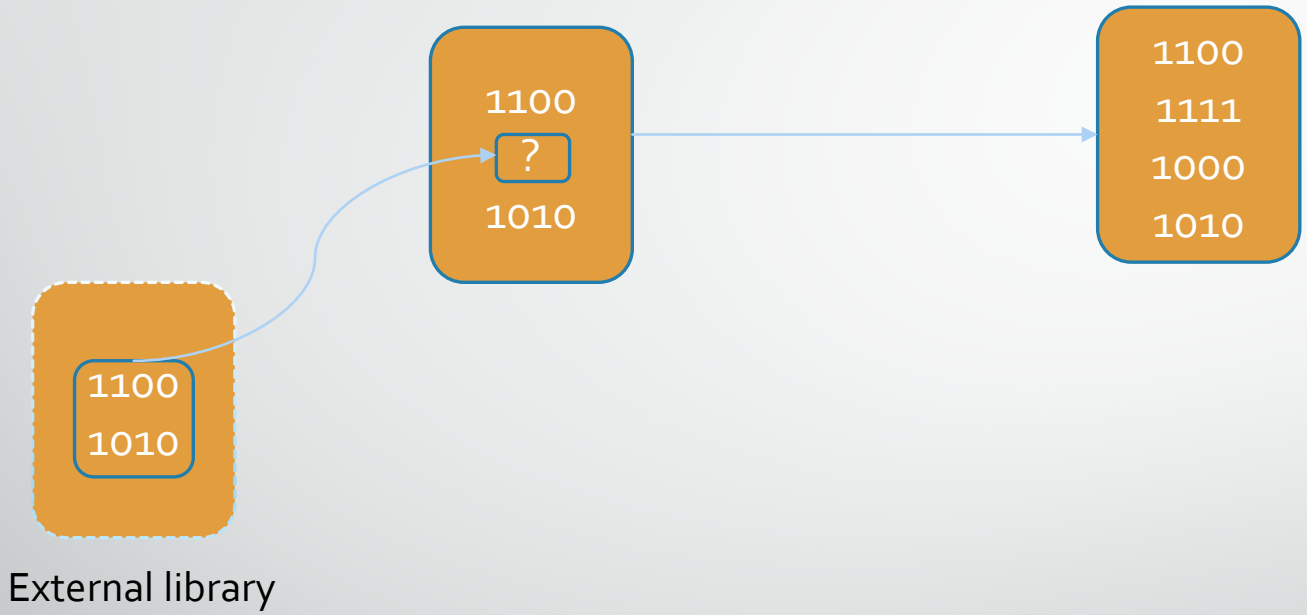
From Assembly language to machine code



The assembler leaves the addresses of external functions undefined

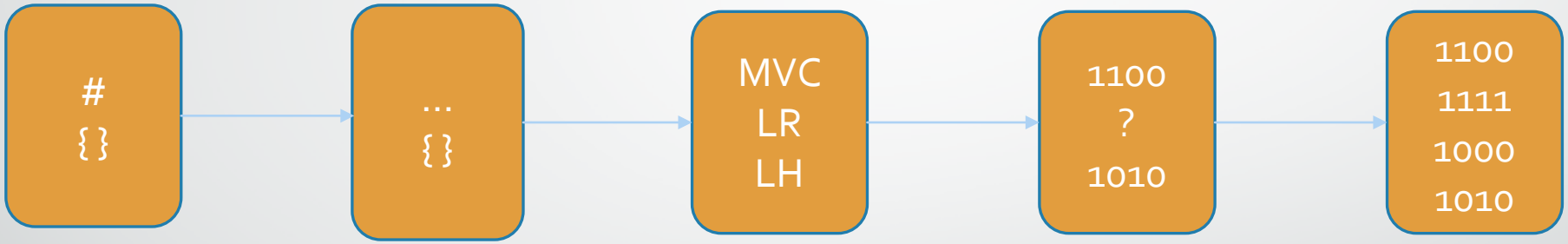
Linking

Create final executable

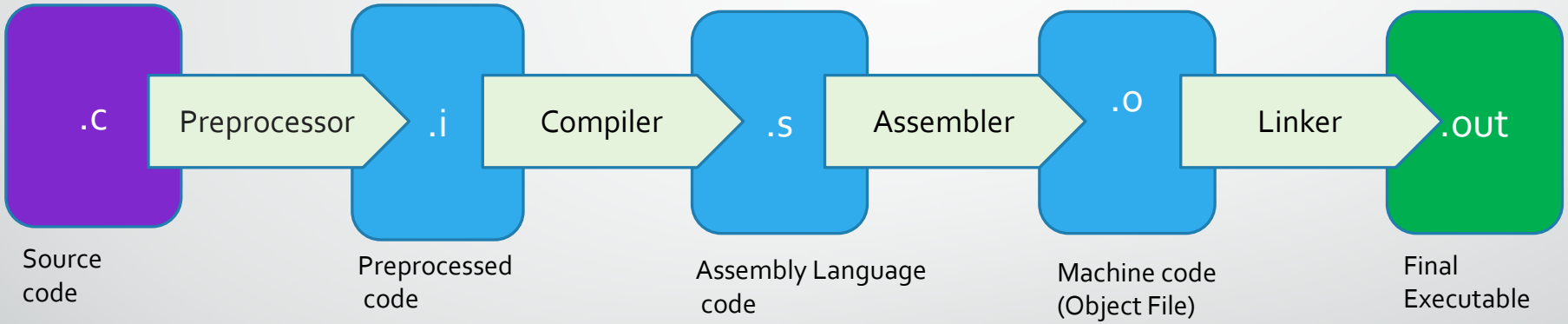


External library

Compilation Stages



Compilation Stages



References

- An Introduction to GCC by Brian Gough
- http://www.hipeac.net/system/files/5_Novillo.pdf
- Compilers: Principles, Techniques, and Tools by Alfred V. Aho, Monica S. Lam, Ravi Sethi and Jeffrey D. Ullman.