

编译原理概述

魏恒峰

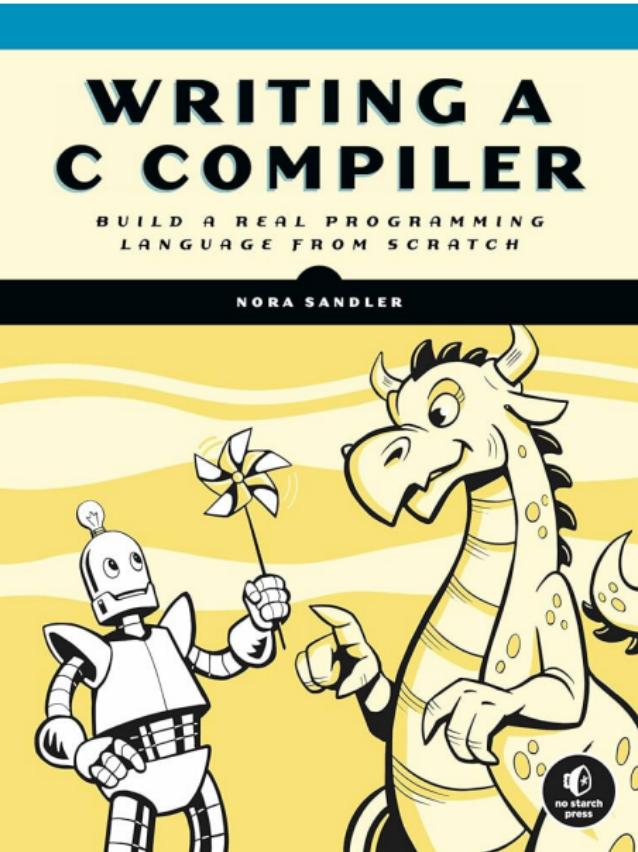
hfwei@nju.edu.cn

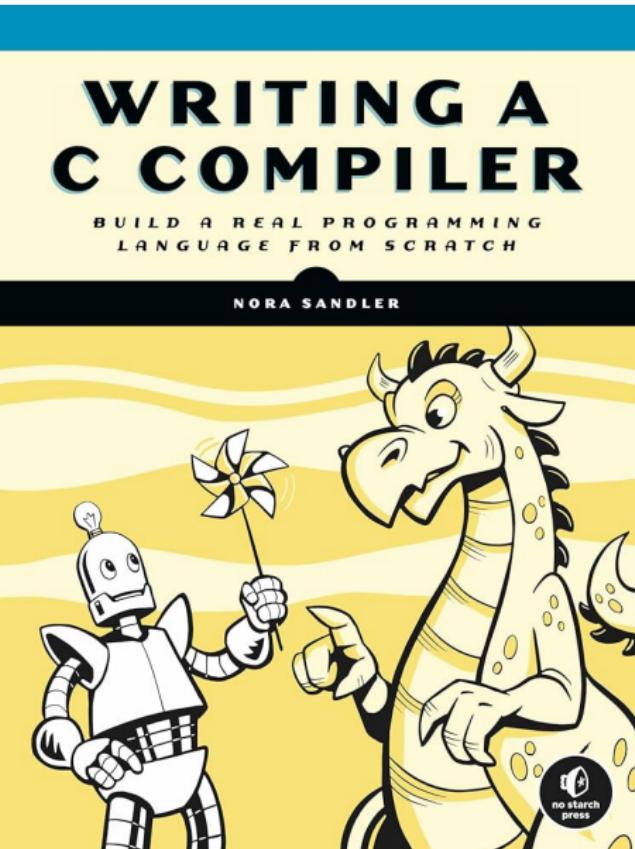
2024 年 03 月 01 日 (周五)





哦？ 编译原理？

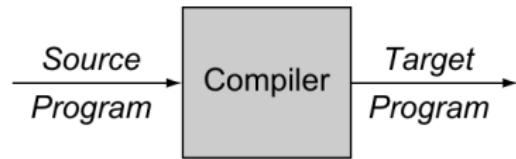




gcc fib.c -o fib

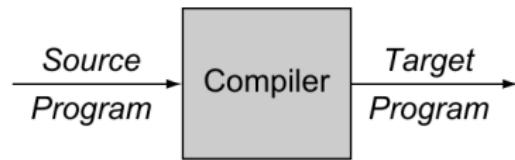
“高级”语言 \Rightarrow (通常) “低级”语言 (如, 汇编语言)

汇编语言经过**汇编器**生成机器语言



“高级”语言 \Rightarrow (通常) “低级”语言 (如, 汇编语言)

汇编语言经过汇编器生成机器语言



GopherJS - A compiler from Go to JavaScript

[godoc](#) [reference](#)

GopherJS compiles Go code ([golang.org](#)) to pure JavaScript code. Its main purpose is to give you the opportunity to write front-end code in Go which will still run in all browsers.

Q : 机器语言是如何跑起来的?

Q : 机器语言是如何跑起来的?

作业 (P1 ~ P9): <https://www.bilibili.com/video/BV1EW411u7th>

(计算机科学速成课 40 集全 Crash Course Computer Science)

Q : 机器语言是如何跑起来的?

作业 (P1 ~ P9): <https://www.bilibili.com/video/BV1EW411u7th>
(计算机科学速成课 40 集全 Crash Course Computer Science)



Q : 机器语言是如何跑起来的?

作业 (P1 ~ P9): <https://www.bilibili.com/video/BV1EW411u7th>
(计算机科学速成课 40 集全 Crash Course Computer Science)



Q : 机器语言是如何跑起来的?

作业 (P1 ~ P9): <https://www.bilibili.com/video/BV1EW411u7th>
(计算机科学速成课 40 集全 Crash Course Computer Science)



“我只想安静地做个码农，你为什么要来强迫我？”



.fun



alda

[install](#) [tutorial](#) [cheat sheet](#) [docs](#) [community](#)

Alda is a text-based programming language for music composition. It allows you to write and play back music using only a text editor and the command line.

```
piano:  
o3  
g8 a b > c d e f+ g | a b > c d e f+ g4  
g8 f+ e d c < b a g | f+ e d c < b a g4  
<< g1/>g/>g/b/>d/g
```

The language's design equally favors aesthetics, flexibility and ease of use.

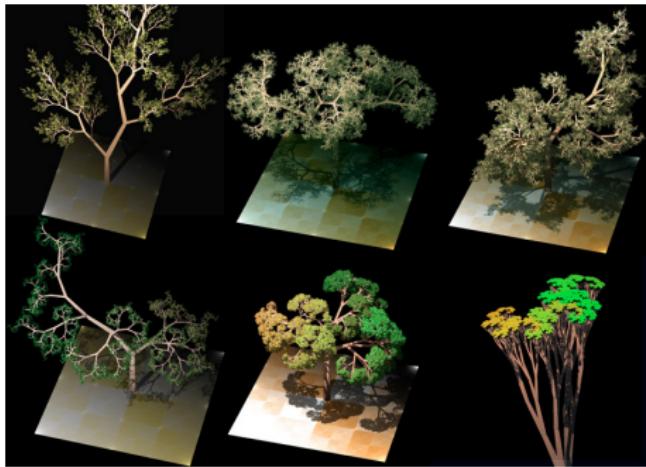
alda repl

alda play -f .alda

alda @ 知乎



10 Demo: Alda @ youtube



Fractal Grower (Try It!)

Sierpinski Triangle @ wiki



语言类应用程序

- ▶ 配置文件解析 (.properties)
- ▶ CSV 文件 (Comma-Separated Values)
- ▶ JSON 文件 (JavaScript Object Notation)

语言类应用程序

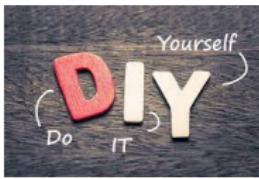
- ▶ 配置文件解析 (.properties)
- ▶ CSV 文件 (Comma-Separated Values)
- ▶ JSON 文件 (JavaScript Object Notation)
- ▶ SQL 引擎 (Structured Query Language)
- ▶ TLA⁺/TLAPS (TPaxos.tla)
- ▶ (Java) 字节码解释器
- ▶ C/C++ 语言编译器

语言类应用程序

- ▶ 配置文件解析 (.properties)
- ▶ CSV 文件 (Comma-Separated Values)
- ▶ JSON 文件 (JavaScript Object Notation)
- ▶ SQL 引擎 (Structured Query Language)
- ▶ TLA⁺/TLAPS (TPaxos.tla)
- ▶ (Java) 字节码解释器
- ▶ C/C++ 语言编译器
- ▶ 排版工具 (LATEX)
- ▶ 绘图工具 (TikZ, Dot/Graphviz)
- ▶ L-System (Cantor Set)

最大的难点：“只见树木，不见森林”





考勤 (00%): 非必要不点名, 不需要请假

平时作业 (00%): ≈ 10 次作业, 每次 1 ~ 2 题

课程实验 (60%): 8 ~ 10 次实验

考勤 (00%): 非必要不点名, 不需要请假

平时作业 (00%): ≈ 10 次作业, 每次 1 ~ 2 题

课程实验 (60%): 8 ~ 10 次实验

期末测试 (40%): 考试周统一安排; 2 小时; **开卷**

每周五发布作业

下周五 23:55 前自愿提交作业



邀请码: 8G928EBJ

课程实验：开发 SysY 语言编译器



小步迭代、多步迭代、贴合课堂讲授进度

L0: 环境配置 下周五 18:00 发布

本科生期末和寒假安排时间表。

周次	星期一	星期二	星期三	星期四	星期五	星期六	星期天
寒假	2月6日 寒假	2月7日 寒假	2月8日 寒假	2月9日 寒假	2月10日 寒假	2月11日 寒假	2月12日 寒假
第1周	2月13日 上课	2月14日 上课	2月15日 上课	2月16日 上课	2月17日 上课	2月18日 考试	2月19日 考试
第2周 停课考试	2月20日 考试	2月21日 考试	2月22日 考试	2月23日 考试	2月24日 考试	2月25日 考试	2月26日 考试
第3周	2月27日 上课	2月28日 上课	3月1日 上课	3月2日 上课	3月3日 上课	3月4日	3月5日
第4周	3月6日 上课	3月7日 上课	3月8日 上课	3月9日 上课	3月10日 上课	3月11日	3月12日
第5周	3月13日 上课	3月14日 上课	3月15日 上课	3月16日 上课	3月17日 上课	3月18日 补缓考	3月19日 补缓考

L1: 词法分析 下下周三 18:00 发布

鼓励讨论，但需独立编码完成课程实验



鼓励讨论，但需独立编码完成课程实验



课程实验：抄袭者当次实验计 0 分

请**经常**使用 Git/GitHub 保留代码编写记录 (OJ 系统原生支持)



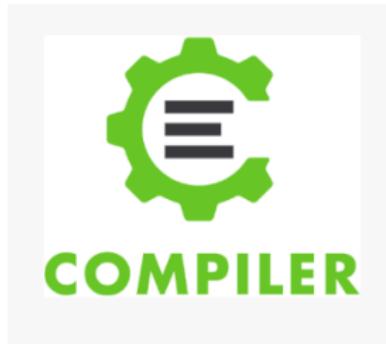
如果你不确定自己的行为**是否构成抄袭**, 请**事先**咨询助教。

QQ 群号: 869910463



助教: 顾龙、李和煦、钱品亦

<http://docs.compilers.cpl.icu/>



搜索

- > 编译原理课程简介
- > 课程要求
- > 教材选购
- > 环境配置
- > LLVM API使用手册
- > 实验

编译原理课程网站，请收藏并及时关注网站更新

courses-at-nju-by-hfwei / **compilers-lectures** Public

Code Issues Pull requests Discussions Actions

master ▾ compilers-lectures / 2023 /

hengxin 2023 README

..

0-overview	+2023/0-overview
README.md	2023 README

<https://github.com/courses-at-nju-by-hfwei/compilers-lectures/tree/master/2023>

 Code

 Issues

 Pull requests

 Actions

 Projects

 Wiki

 master 

 1 branch

 0 tags

 Go to file

 Add file



hengxin +stringtemplate, +bison-ag

 attributed-grammars +stringtemplate, +bison-ag

 cfg +papers

 code-generation +stringtemplate, +bison-ag

 compilers +papers for parsers

 history +papers for parsers

 ir +papers for parsers

 optimization +papers for parsers

 parsing +papers for parsing

 programming-languages +paper: FunctionalStyle

 re-nfa-dfa +papers

 teaching +papers for parsers

The tomassetti.me website has changed: it is now part of strumenta.com. You will continue to find all the news with the usual quality, but in a new layout.

Our articles

Here we talk about parsing, tools and libraries, natural language processing and Kotlin

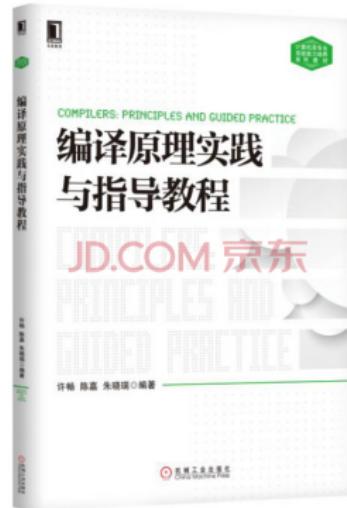
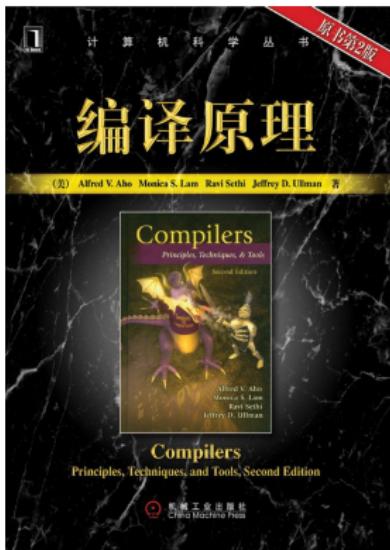
Edited by Federico Tomassetti



tomassetti.me

专业人士，专业网站，推荐订阅

(本书仅供参考)



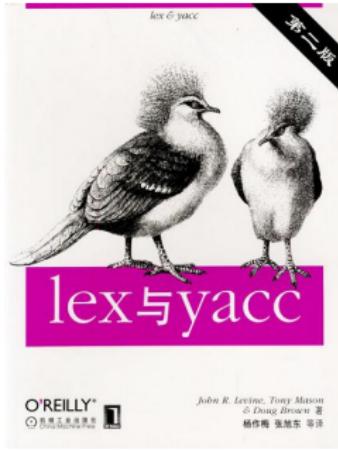
也可使用“**本科教学版**”

[https://cs.nju.edu.cn/
changxu/2_compiler/index.html](https://cs.nju.edu.cn/changxu/2_compiler/index.html)

(已对课程实验做了大幅改动)



Flex: 词法分析器生成器



Bison: 语法分析器生成器

不够现代, 本学期课程实验**不再支持**这些工具



(Since 1988)

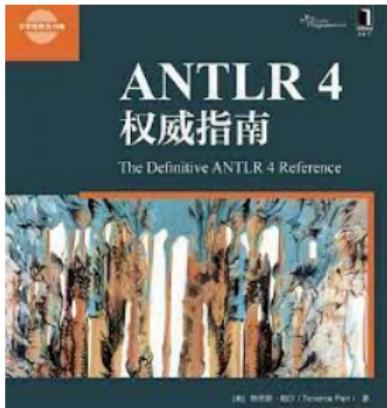
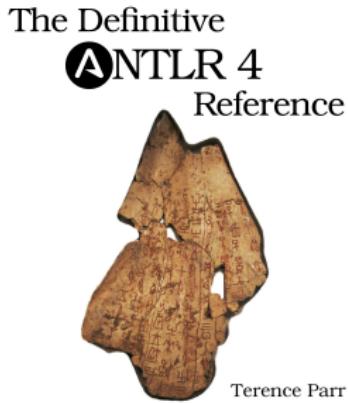


Terence Parr (University of San Francisco)

<https://www.antlr.org/index.html>

<https://www.antlr.org/tools.html> (IntelliJ Plugin)

<http://lab.antlr.org/> (Online lab)



基于 ANTLR 4，是课程实验指导的**重要**参考资料

Language Implementation Patterns

Create Your Own Domain-Specific and General Programming Languages

Terence Parr



Language
Implementation Patterns

编程语言 实现模式

Create Your Own Domain-Specific
and General Programming Languages

[译] Terence Parr 著
李直来 译
高鹏程 审校



基于 ANTLR 3，与 ANTLR 4 相比，有些过时，
但可以看作理解 ANTLR 4 的基础

自制编译器

How to Develop a Compiler

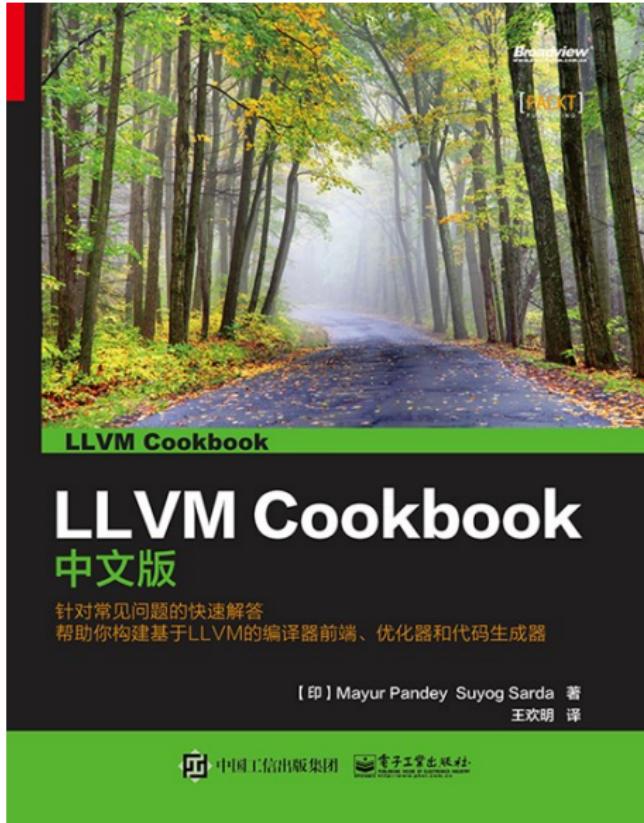
[日] 青木峰郎 / 著 严基连 焦云 / 译

Turing
图灵精英
设计丛书

从零开始制作真正的编译器
贯穿编译、汇编、链接、加载的全过程！
比“是书”更具实践性！

中国工信出版集团 人民邮电出版社
POSTS & TELECOM PRESS

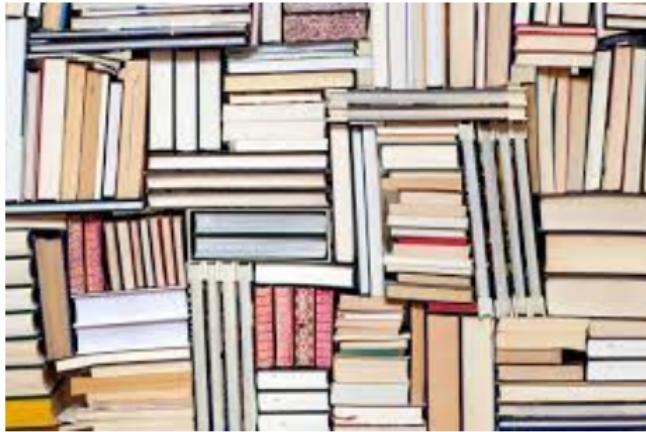
“从零开始制作真正的编译器”，对课程实验很有帮助，**强烈推荐**



从某次实验开始，你就会开始接触 LLVM (<https://llvm.org/>)



LLVM 简介 @ Bilibili



更多参考书, 随课程进展陆续发布

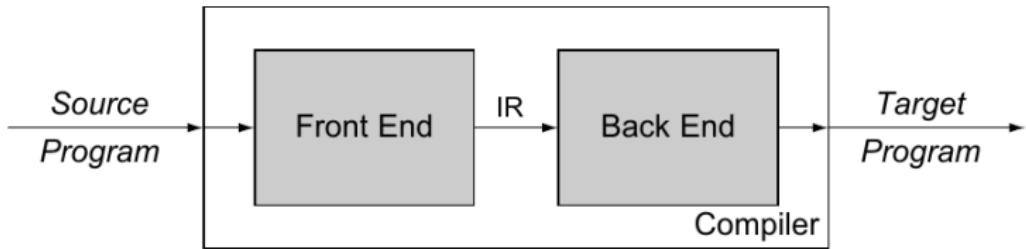
<http://47.122.3.40:8081>

LET'S GET
STARTED



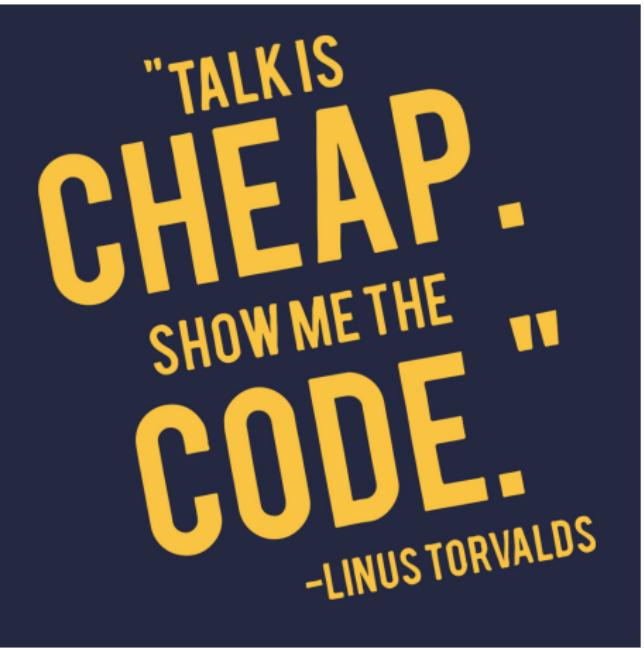


IR: Intermediate Representation (中间表示)



前端（分析阶段）: 分析源语言程序, 收集所有必要的信息

后端（综合阶段）: 利用收集到的信息, 生成目标语言程序

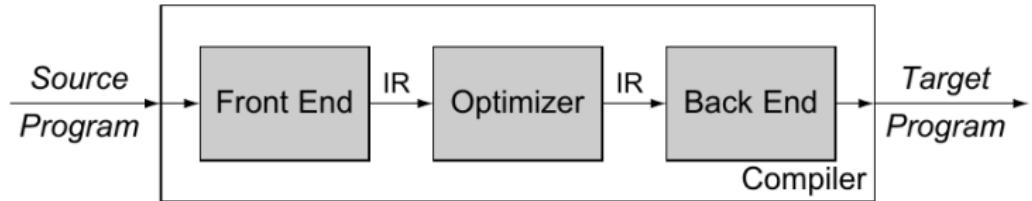


Clang: a C language family frontend for LLVM

The Clang project provides a language front-end and tooling infrastructure for languages in the C language family (C, C++, Objective C/C++, OpenCL, CUDA, and RenderScript) for the [LLVM](#) project. Both a GCC-compatible compiler driver (`clang`) and an MSVC-compatible compiler driver (`clang-cl.exe`) are provided. You can [get and build](#) the source today.

<https://clang.llvm.org/>

(`factorial.c`)



机器无关的中间表示优化

Clang: a C language family frontend for LLVM

The Clang project provides a language front-end and tooling infrastructure for languages in the C language family (C, C++, Objective C/C++, OpenCL, CUDA, and RenderScript) for the [LLVM](#) project. Both a GCC-compatible compiler driver (`clang`) and an MSVC-compatible compiler driver (`clang-cl.exe`) are provided. You can [get and build](#) the source today.

<https://clang.llvm.org/>

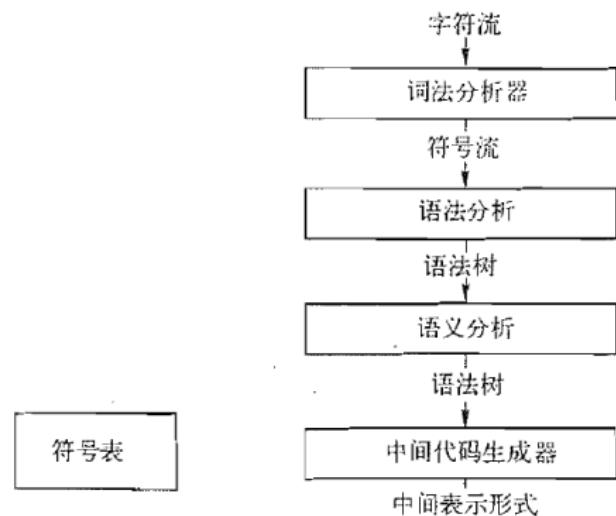
(`opt.c`)

在设计实际生产环境中的编译器时，**优化**通常占用了大多数时间

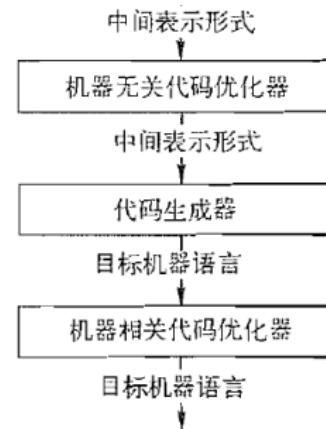


Maple IR 及其各种优化技术

编译器前端：分析阶段



编译器后端：综合阶段



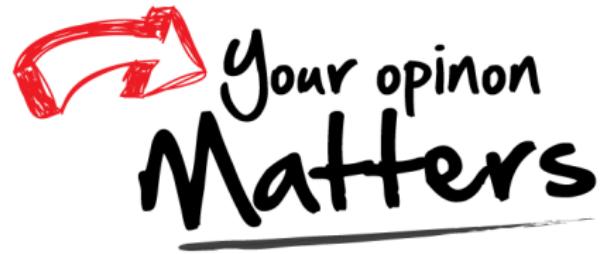
Clang: a C language family frontend for LLVM

The Clang project provides a language front-end and tooling infrastructure for languages in the C language family (C, C++, Objective C/C++, OpenCL, CUDA, and RenderScript) for the [LLVM](#) project. Both a GCC-compatible compiler driver (`clang`) and an MSVC-compatible compiler driver (`clang-cl.exe`) are provided. You can [get and build](#) the source today.

<https://clang.llvm.org/>

(`naming.c`)

Thank You!



Office 926

hfwei@nju.edu.cn