编译原理

1、编译与编译器

编译

把一门语言编写的程序(源程序)翻译成另外一门语言程序(目标代码)。

编译器

完成编译工作的程序。

2、编译流程

程序的编译是一个系统性的知识(也是大学计算机专业的主课程)。一个编译器通常包含如下步骤:



2-1、词法分析

词法分析是编译的第一步,主要任务是读入源程序的输入字符,将其组成词素(Token: 单词、符号等元素)并标记词素类型(如字符串、数字、运算符、关键字等)和词素位置(行列),这个阶段也包含一些预处理任务,如过滤合并空格、注释……等。

有限状态自动机 (FSM/FSA)

研究事物解决能力的一种有效模型 zh.wikipedia.org https://zh.wikipedia.org/wiki/有限状态机 - 通信协议

- 词法分析

2-2、语法分析

语法分析是根据某种给定的 形式文法 对词法分析阶段得到的 Token 进行分析并确定其语法结构的过程。

形式文法指的是字符串序列在某种语言中文法组织规则,如自然语言中的肯定句、疑问句,程序中的 条件、循环、赋值等。





巴科斯范式

一种用来描述(表示)某种语言文法的范式(符号集),广泛应用于各种程序设计语言。

产生式: 文法格式, 推导规则(非终结符的分解规则)。

非终结符:可再分解推导的内容。

终结符: 不可再分解推导的内容。

zh.wikipedia.org

https://zh.wikipedia.org/wiki/巴科斯范式

BNF for EcmaScript.jj

tomcopeland.blogs.com

JavaScript: The Good Parts

Appendix D. Syntax Diagrams Thou map of woe, that thus dost talk in signs! —William Shakespeare, The Tragedy of Titus Andronicus - Selection from JavaScript: The Good Parts [Book]

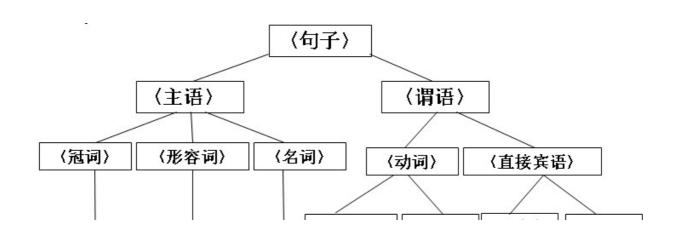
o www.oreilly.com



JSON



www.json.org



当定义了语言的文法后,如何去分析给定字符串内容是否符合该语言定义的文法就是后续 文法解析 要考虑的事情了。

LL 分析器

zh.wikipedia.org

https://zh.wikipedia.org/wiki/LL剖析器

LL(1)分析_百度百科

LL(1)分析使用显式栈而不是递归调用来完成分析。以标准方式表示这个栈非常有用,这样LL(1)分析程序的动作就可以快捷地显现出 来。



AST

AST explorer

An online AST explorer

3、手动实现一个编译器

3-1、语法分析器

课件代码.....

3-2、词法分析器

课件代码.....



编译原理相关的前端应用

- vue-template, react-jsx
- markdown
- webpack: markdown-loader