



编译系统

第一章 绪论

哈尔滨工业大学 陈鄞



第1讲（绪论）要点

➤ 什么是编译

➤ 将高级语言翻译成汇编语言或机器语言的过程

➤ 编译系统的组成及各部分的主要任务

➤ 词法分析

➤ 语法分析

➤ 语义分析

➤ 中间代码生成

➤ 目标代码生成

➤ 代码优化

第1讲（绪论）要点

➤ 什么是编译

➤ 将高级语言翻译成汇编语言或机器语言的过程

➤ 编译系统的组成及各部分的主要任务

➤ 词法分析 识别单词，确定单词的类型

➤ 语法分析

➤ 语义分析

➤ 中间代码生成

➤ 目标代码生成

➤ 代码优化

token: <种别码, 属性值>

↑

词法单元

关键字
一词一码
标识符
多词一码
常量
一型一码
运算符
一词一码
界符
一词一码

第1讲（绪论）要点

➤ 什么是编译

➤ 将高级语言翻译成汇编语言或机器语言的过程

➤ 编译系统的组成及各部分的主要任务

➤ 词法分析 识别单词，确定单词的类型

➤ 语法分析 识别短语，构造语法分析树

➤ 语义分析 { (声明语句) 收集标识符的属性信息

➤ 中间代码生成

➤ 目标代码生成

➤ 代码优化

种属 (kind)
类型 (type)
地址
长度
作用域
参数与返回值信息
.....



符号表

第1讲（绪论）要点

➤ 什么是编译

➤ 将高级语言翻译成汇编语言或机器语言的过程

➤ 编译系统的组成及各部分的主要任务

➤ 词法分析 识别单词，确定单词的类型

➤ 语法分析 识别短语，构造语法分析树

➤ 语义分析 { (声明语句) 收集标识符的属性信息
语义检查

➤ 中间代码生成

➤ 目标代码生成

➤ 代码优化

变量或过程名重复声明

变量或过程未经声明就使用

运算分量类型不匹配

操作符与操作数之间的类型不匹配

对非数组变量使用数组访问操作符
数组下标不是整数

对非过程名使用过程调用操作符
过程调用的参数类型或数目不匹配
函数返回类型有误

$a[x]$

$a(...)$

常用的三地址指令

序号	指令类型	指令形式
1	赋值指令	$x = y \text{ op } z$ $x = \text{op } y$
2	复制指令	$x = y$
3	条件跳转	if $x \text{ relop } y$ goto n
4	非条件跳转	goto n
5	参数传递	param x
6	过程调用 函数调用	call p, n $y = \text{call } p, n$
7	过程返回	return x
8	数组引用	$x = y[i]$
9	数组赋值	$x[i] = y$
10	地址及 指针操作	$x = \& y$ $x = *y$ $*x = y$

地址可以具有如下形式之一

- 源程序中的名字 (*name*)
- 常量 (*constant*)
- 编译器生成的临时变量 (*temporary*)

三地址指令的四元式表示

- $x = y \text{ op } z$ (**op** , y , z , x)
- $x = \text{op } y$ (**op** , y , $_$, x)
- $x = y$ (**=** , y , $_$, x)
- if $x \text{ relop } y$ goto n (**relop** , x , y , n)
- goto n (**goto** , $_$, $_$, n)
- param x (**param** , $_$, $_$, x)
- $y = \text{call } p, n$ (**call** , p , n , y)
- return x (**return** , $_$, $_$, x)
- $x = y[i]$ (**=[]** , y , i , x)
- $x[i] = y$ (**[]=** , y , x , i)
- $x = \&y$ (**&** , y , $_$, x)
- $x = *y$ (**=*** , y , $_$, x)
- $*x = y$ (***=** , y , $_$, x)



三地址指令序列唯一确定了运算完成的顺序

第1讲（绪论）要点

➤ 什么是编译

➤ 将高级语言翻译成汇编语言或机器语言的过程

➤ 编译系统的组成及各部分的主要任务

➤ 词法分析 识别单词，确定单词的类型

➤ 语法分析 识别短语，构造语法分析树

➤ 语义分析 { (声明语句) 收集标识符的属性信息
语义检查

➤ 中间代码生成

➤ 目标代码生成

➤ 代码优化 运行更快或占用空间更少



结束

