Atitit 词法分析的常见问题

目录

[1.1. 如何区别关键字和标识符呢？ 2 1](#_Toc10009)

[13、 error 出错处理的函数，处理出现的词法错误。、 1](#_Toc355)

[1.2. 区分标识符和常数 1](#_Toc18229)

[1.3. 词法分析的逻辑结果包括空格吗 好像不包括， 2](#_Toc8280)

[1.4. 分号逗号是包括的。 2](#_Toc23317)

[1.5. 词法分析的逻辑结果包括分界符号码 看起来好像包括 2](#_Toc18076)

[1.6. 词法分析的第一阶段即扫描器 第二阶段评估器（Evaluator） 2](#_Toc9788)

## 如何区别关键字和标识符呢？ 2

如何区别关键字和标识符呢？

因为在很多语言中关键字和标识符是有交集的，但从词法分析上看，关键字是标识符的一部分

第一种，我们需要将上面的转移图稍加修改：（以识别if为例）

　　　　此处还应注意一个问题，假如输入ifxy呢？并不是关键字，因此这里还需判断。

再看第二种，使用关键字表，即Hash表。对语言中所有的关键字构建哈希表，对所有的关键字和标识符，先统一按标识符的转移表进行识别，识别完成后，再查哈希表看是否为关键字。通过构造合理的哈希表，可以再O(1)时间内完成。

## error 出错处理的函数，处理出现的词法错误。、

1. 有一类词法错误可以在词法分析时发现，如出现字母表以外的非法字符、不合规则的常数、标识符的前缀为保留字等。但还有一类词法错误，例如，把if写成fi，词法分析会将fi当作标识符处理，le ngth中间多了一个空格，词法分析会将le和ngth当作两个标识符处理。这类词法错误往往要推迟到语法分析时才能发现，不属于函数处理的范畴。

## 区分标识符和常数

普通source有间隔符双引号区分。。。Cmd 里面的话，更具cur state来区分

Qa

## 词法分析的逻辑结果包括空格吗 好像不包括，

说是评估器把它抑制了，扫描器是包括的。

参考 atitit.c词法分析器的详细原理.docx。。参考请问编译原理词法分析器是干什么用的，怎么用？？？-CSDN论坛-CSDN.NET-中国最大的IT技术社区.html， 也是不包括

## 分号逗号是包括的。

说是评估器把它抑制了，扫描器是包括的。

## 词法分析的逻辑结果包括分界符号码 看起来好像包括

说是评估器把它抑制了，扫描器是包括的。

## 词法分析的第一阶段即扫描器 第二阶段评估器（Evaluator）

词法分析的第一阶段即扫描器，通常基于有限状态自动机

词法分析的第二阶段评估器（Evaluator）

，语法分析器需要第二阶段的评估器（Evaluator）。评估器根据语素中的字符序列生成一个“值”，这个“值”和语素的类型便构成了可以送入语法分析器的单词。一些诸如括号的语素并没有“值”，评估器函数便可以什么都不返回。整数、标识符、字符串的评估器则要复杂的多。评估器有时会抑制语素，被抑制的语素（例如空白语素和注释语素）随后不会被送入语法分析器。