

Chapter-2-词法分析2

main topic: 手写词法分析器

Course Note

常见表达

- $digit = [0 - 9]$
- $digits = digit^+ = [1 - 9]$
- $number = digits(.digits)?(E[+-]?digits)?$
 - 正整数、小数点、指数
- $letter = [A - Za - z]$
- $id = letter(letter|digit)^*$
- $relop = <> | <= | >= | <>$

手写的技巧

- 向前看、向前走、调整状态
- 记录来时最长匹配、无路可走便回头（回溯到上一个正确位置）
 - `nextToken()`
 - `while (nextToken())`

如何区分 int、real 与 sci?

num: 整数部分[. 可选的小数部分][E[可选的 +-] 可选的指数部分]

- 在 real 与 sci 中, 有时需要回退, 寻找最长匹配
- 根据下一个字符即可判定词法单元的类型; 否则, 调用错误处理模块 (对应 other), 报告该字符有误, 忽略该字符

Slide

[2-lexer-handwritten](#)