



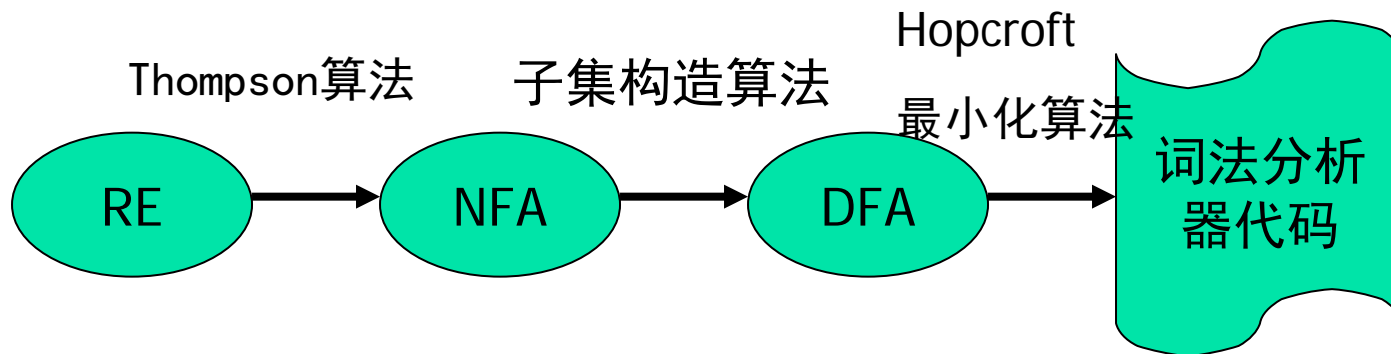
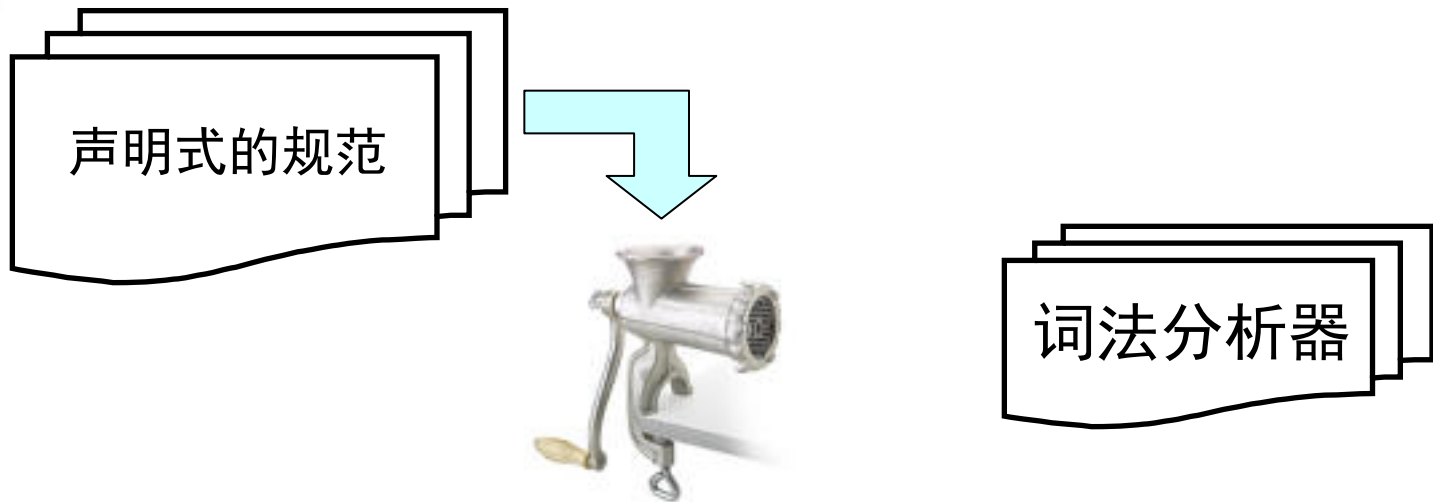
词法分析---正则表达式到非确定有限状态自动机

编译原理

华保健

bjhua@ustc.edu.cn

回顾：自动生成



RE -> NFA:

Thompson算法

- 基于对RE的结构做归纳
 - 对基本的RE直接构造
 - 对复合的RE递归构造
- 递归算法，容易实现
 - 在我们的实现里，不到100行的C代码

RE \rightarrow NFA:

Thompson algorithm

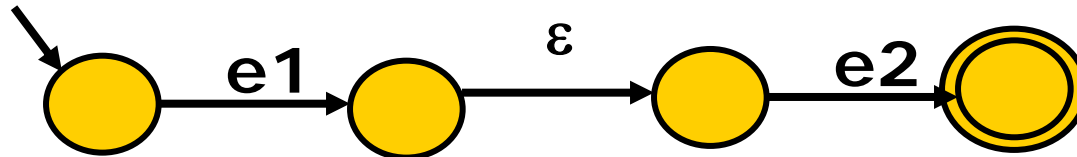
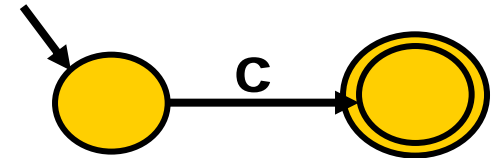
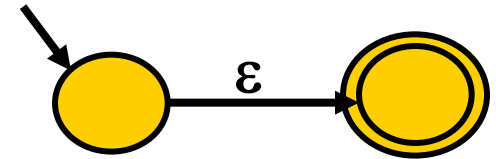
$e \rightarrow \epsilon$

$\rightarrow c$

$\rightarrow e_1 e_2$

$\rightarrow e_1 \mid e_2$

$\rightarrow e_1^*$



RE \rightarrow NFA:

Thompson algorithm

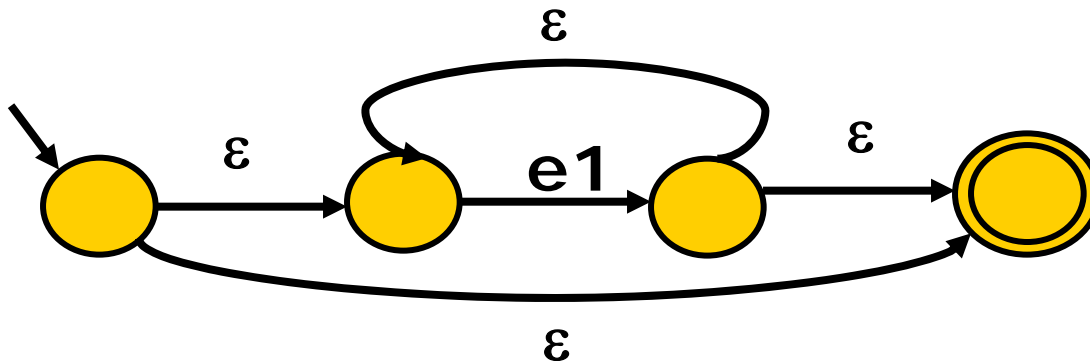
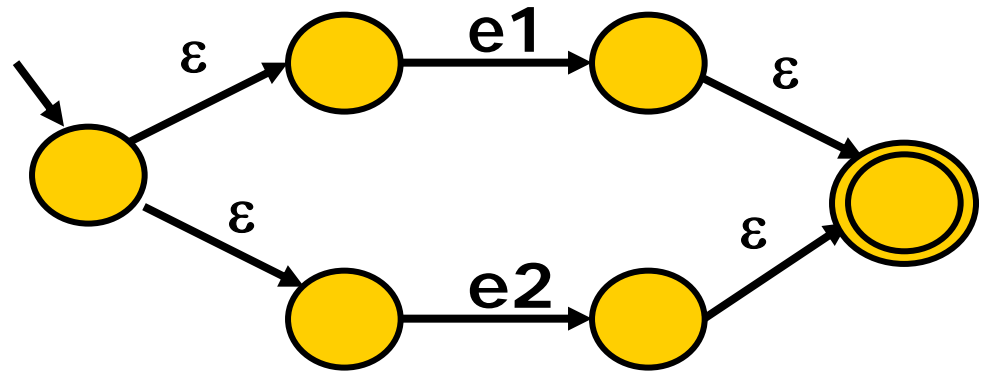
$e \rightarrow \epsilon$

$\rightarrow c$

$\rightarrow e1\ e2$

$\rightarrow e1 \mid e2$

$\rightarrow e1^*$





示例

$a(b|c)^*$