



编译技术 词法分析

大连理工大学软件学院

正规式 \Rightarrow NFA \Rightarrow DFA

途径3

先构造NFA，再将NFA转换为DFA

三大步骤：

- NFA构建
- NFA \rightarrow DFA的转化（子集构造法）
- DFA化简

NFA

\Rightarrow

DFA

理论依据

根据有限自动机理论，设L为一个有不确定的有限自动机接受的集合，则存在一个接受L的确定的有限自动机

怎样进行NFA到DFA的转化

子集构造法

子集构造法

NFA

\Rightarrow

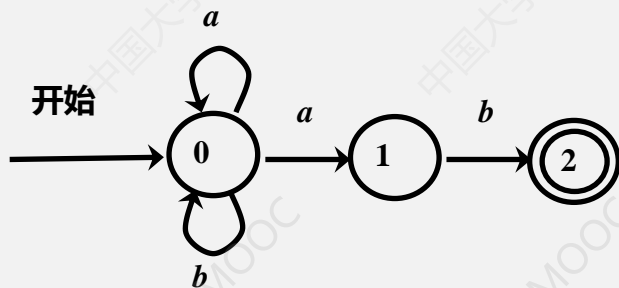
DFA

DFA的一个状态是NFA的一个状态集合

读了输入 $a_1 a_2 \dots a_n$ 后

NFA能到达的所有状态: s_1, s_2, \dots, s_k ,
则DFA到达状态 $\{s_1, s_2, \dots, s_k\}$

识别语言 $(a|b)^*ab$ 的DFA



子集构造法

NFA

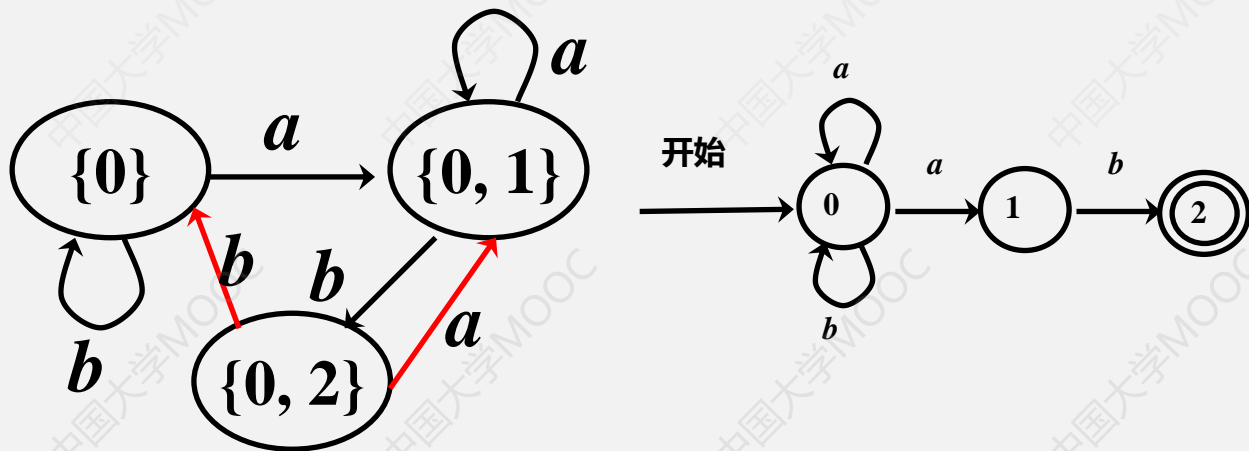
\Rightarrow

DFA

DFA的一个状态是NFA的一个状态集合

读了输入 $a_1 a_2 \dots a_n$ 后

NFA能到达的所有状态: s_1, s_2, \dots, s_k ,
则DFA到达状态 $\{s_1, s_2, \dots, s_k\}$



2.3 有限自动机

01

$A = \{0, 1, 2, 4, 7\}$

02

$B = \{1, 2, 3, 4, 6, 7, 8\}$

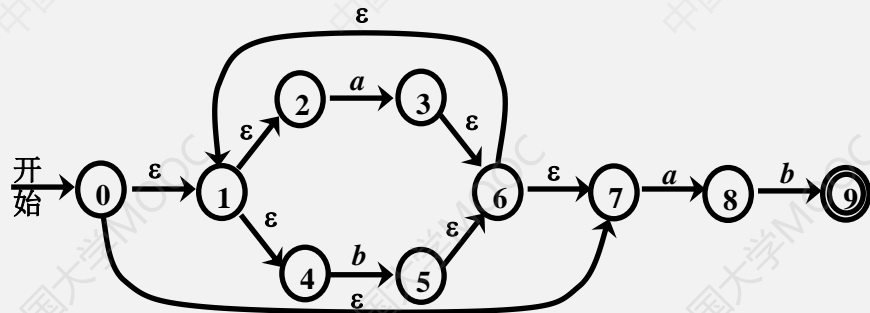
03

$C = \{1, 2, 4, 5, 6, 7\}$

04

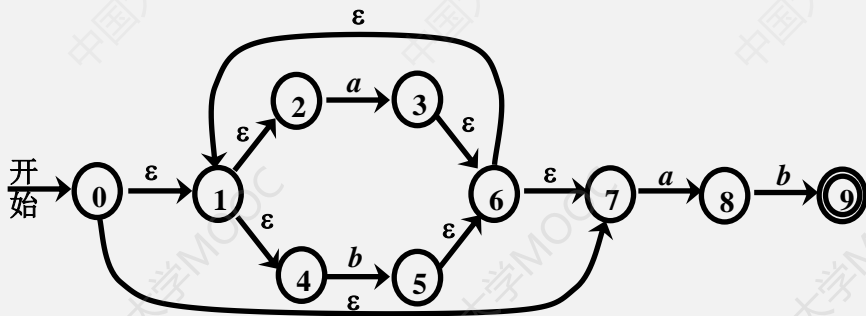
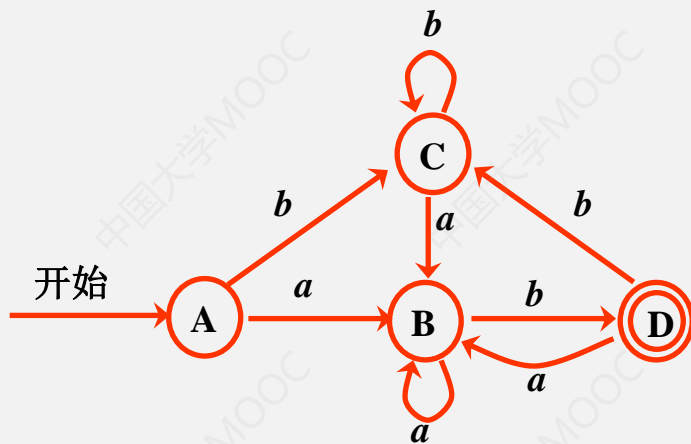
$D = \{1, 2, 4, 5, 6, 7, 9\}$

状态	输入符号	
	<i>a</i>	<i>b</i>
<i>A</i>	<i>B</i>	<i>C</i>
<i>B</i>	<i>B</i>	<i>D</i>
<i>C</i>	<i>B</i>	<i>C</i>
<i>D</i>	<i>B</i>	<i>C</i>



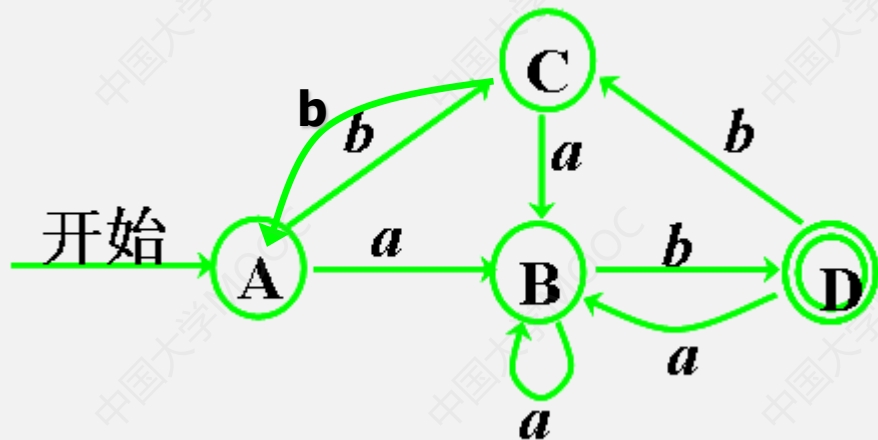
2.3 有限自动机

状态	输入符号	
	<i>a</i>	<i>b</i>
<i>A</i>	<i>B</i>	<i>C</i>
<i>B</i>	<i>B</i>	<i>D</i>
<i>C</i>	<i>B</i>	<i>C</i>
<i>D</i>	<i>B</i>	<i>C</i>



2.3 有限自动机

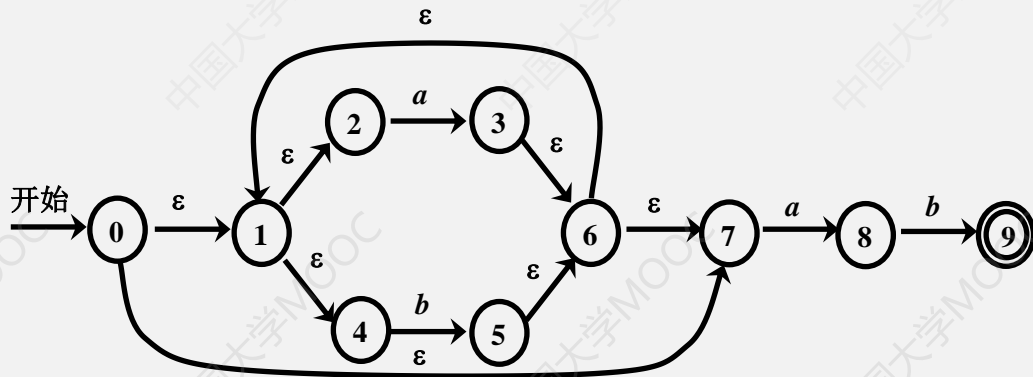
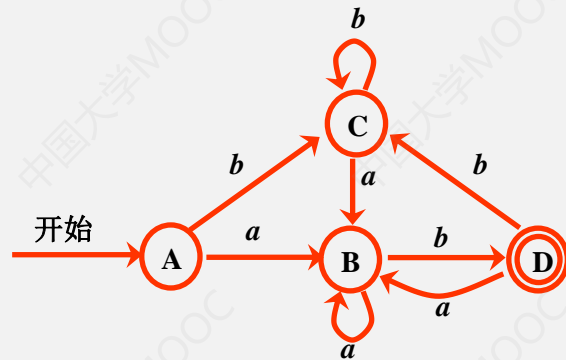
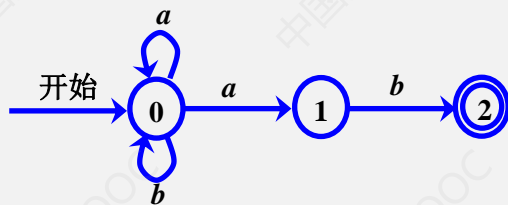
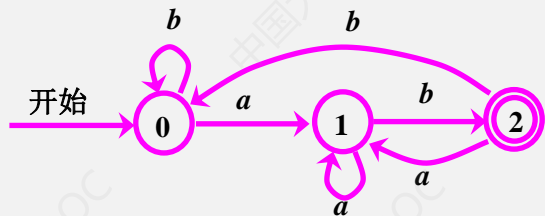
状态	输入符号	
	<i>a</i>	<i>b</i>
<i>A</i>	<i>B</i>	<i>C</i>
<i>B</i>	<i>B</i>	<i>D</i>
<i>C</i>	<i>B</i>	<i>A</i>
<i>D</i>	<i>B</i>	<i>C</i>



2.3 有限自动机



识别语言 $(a|b)^*ab$ 的自动机



正规式 \Rightarrow NFA \Rightarrow DFA

途径3

先构造NFA，再将NFA转换为DFA

三大步骤：

- NFA构建
- NFA \rightarrow DFA的转化（子集构造法）
- **DFA化简**

2.3 有限自动机



DFA的化简



通过算法构造的NFA，而后经过子集构造法得来的DFA通常不是最简的。



如何判定一个DFA是不是最简，that's a question



DFA 的 化简



判断依据

状态的可区分性



状态的可区分性

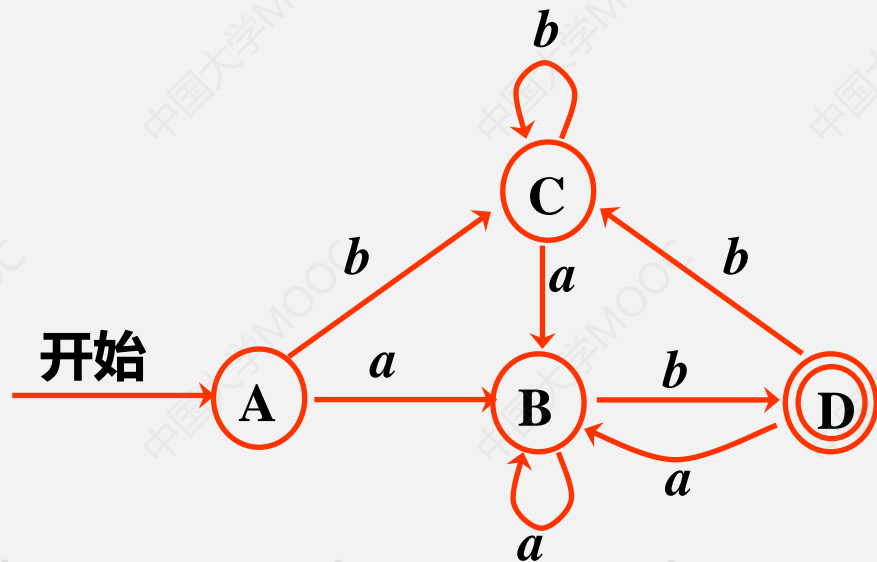
存在串 w ，使得从状态 s 开始，对 w 进行状态转换，最终停在某个接受状态；而对于从 t 开始对 w 进行状态转换后，却停在某个非接受状态

2.3 有限自动机

可区别的状态

A和B是可区别的状态

A和C是不可区别的状态



DFA 的 化简

DFA化简的途径

根据状态是否可以区分，将状态划分成若干个集合，每个集合内的状态之间都不可区分，而任意两个集合中的元素都是可以互相区分的。

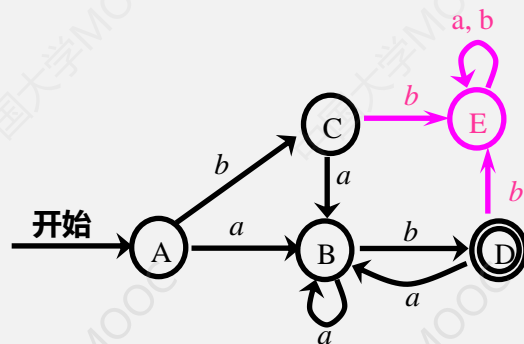
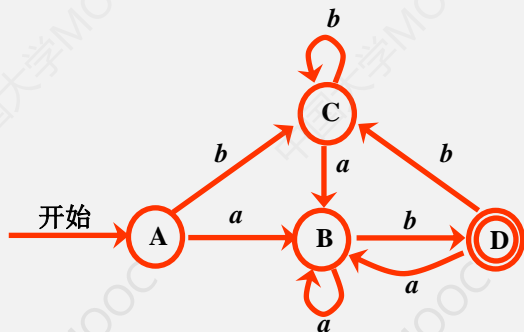
依据原始的DFA，在合并后的状态基础上，建立新的状态转换关系

2.3 有限自动机



化简时DFA的状态转换函数必须是一个**全函数**

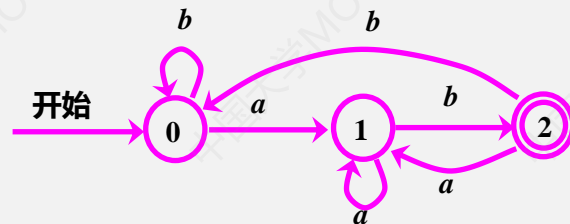
- 死状态 ➤ 左图无须加死状态 ➤ 右图加入死状态E



方法

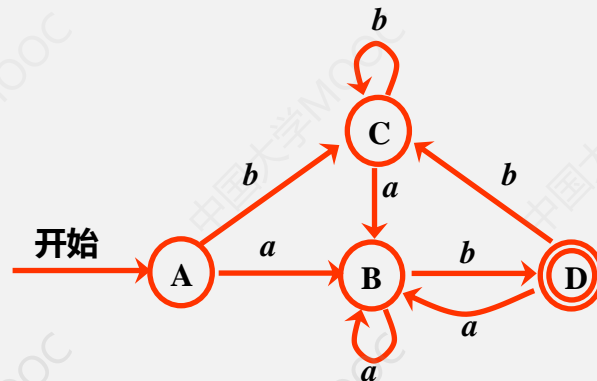
01

$\{A, B, C\}, \{D\}$
 $move(\{A, B, C\}, a) = \{B\}$
 $move(\{A, \textcolor{red}{B}, C\}, b) = \{C, \textcolor{red}{D}\}$



02

$\{A, C\}, \{B\}, \{D\}$
 $move(\{A, C\}, a) = \{B\}$
 $move(\{A, C\}, b) = \{C\}$





编译技术 词法分析

大连理工大学软件学院