

第三讲

非确定有限自动机

2022/3/8

School of Software

1

非确定有限自动机

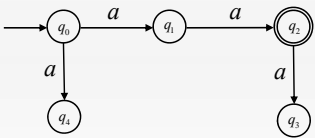
- 非确定有限自动机的概念
- ϵ -转移
- 非确定有限自动机的定义
- 扩展转移函数
- 等价性证明
- 文本搜索

非确定有限自动机

- 非确定有限自动机的概念
- ϵ -转移
- 非确定有限自动机的定义
- 扩展转移函数
- 等价性证明
- 文本搜索

非确定有限自动机的概念

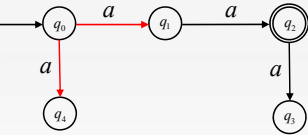
$\Sigma=\{a\}$



非确定有限自动机的概念

$\Sigma=\{a\}$

两个
变迁选择

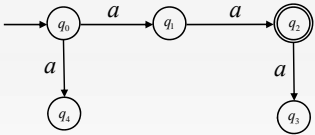


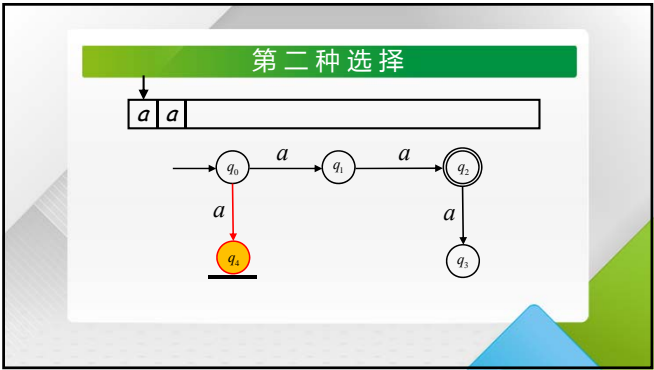
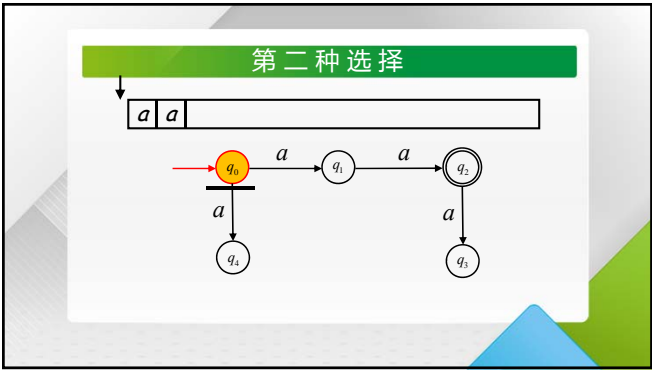
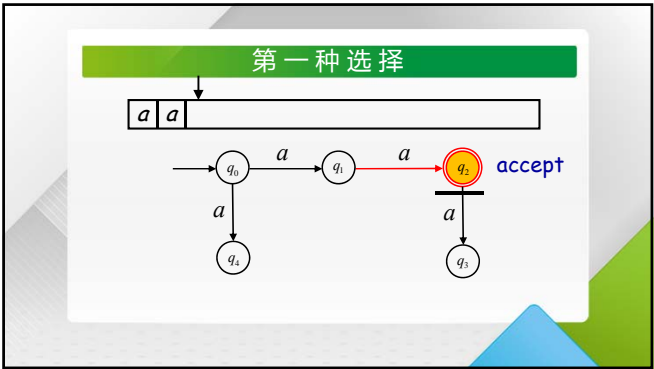
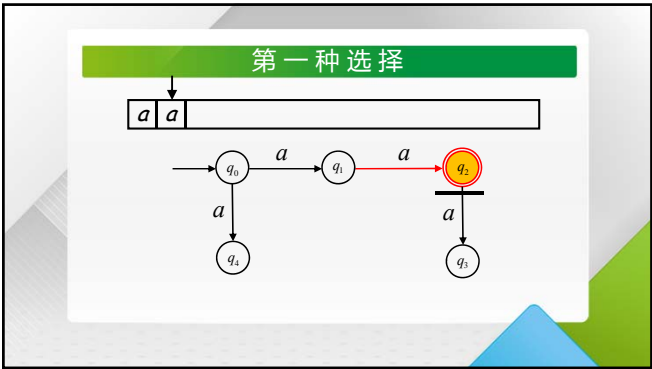
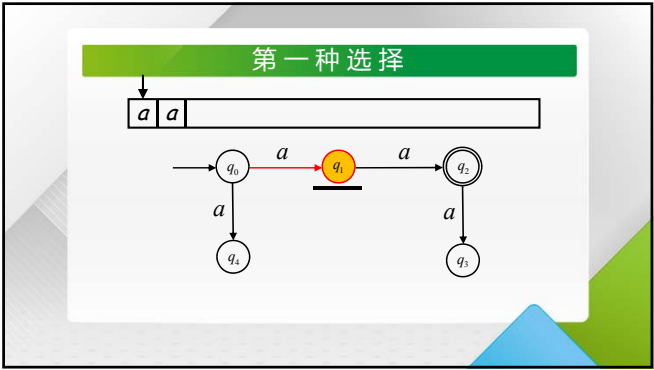
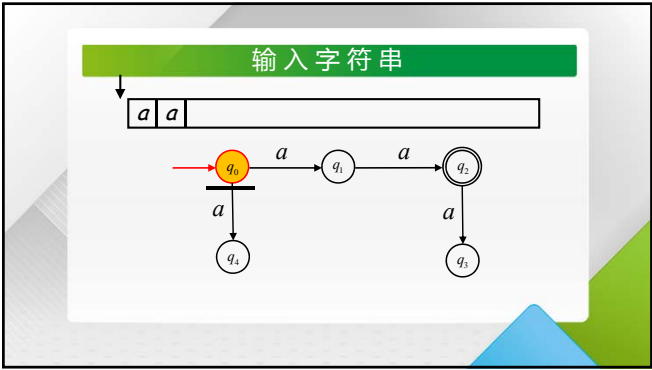
非确定有限自动机的概念

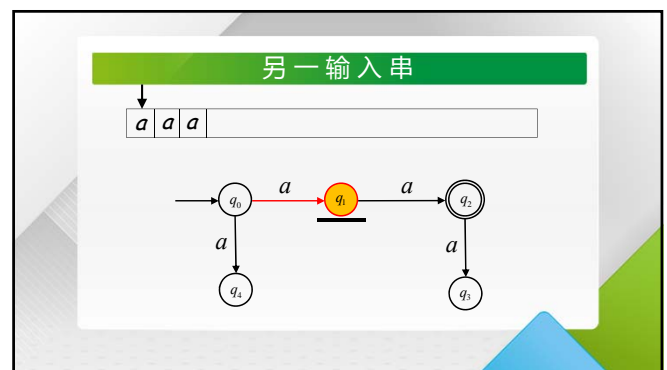
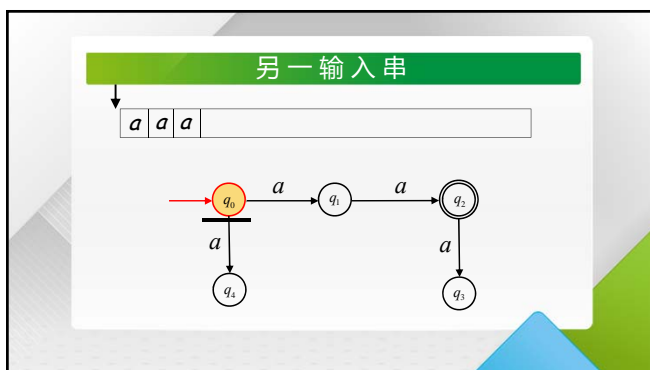
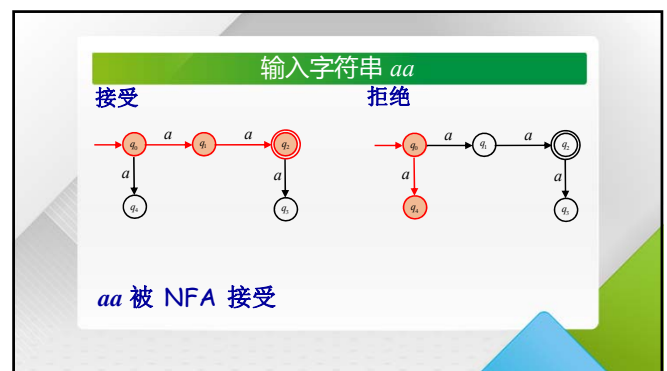
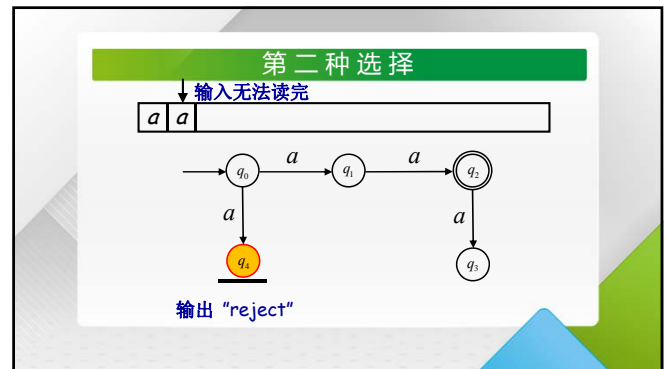
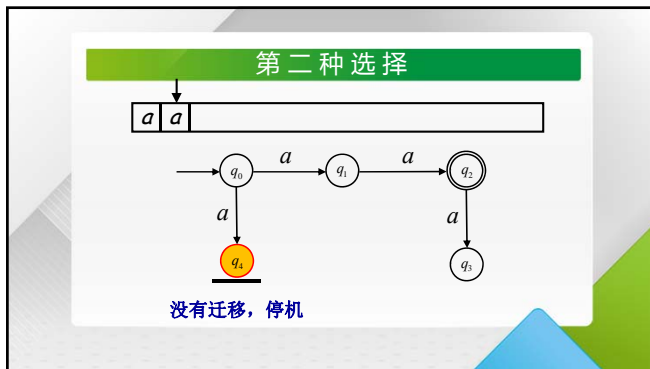
$\Sigma=\{a\}$

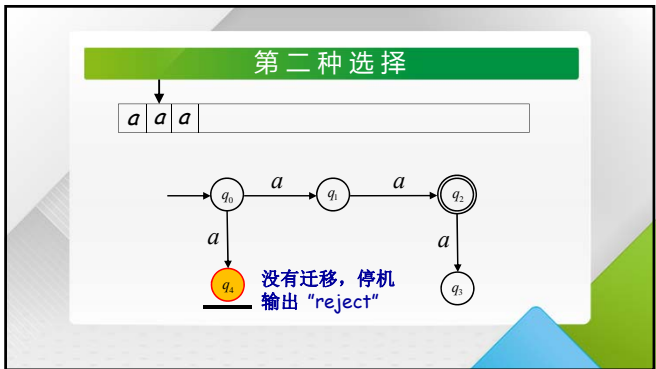
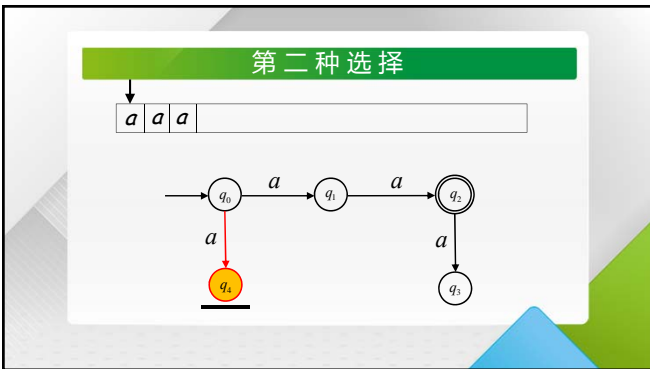
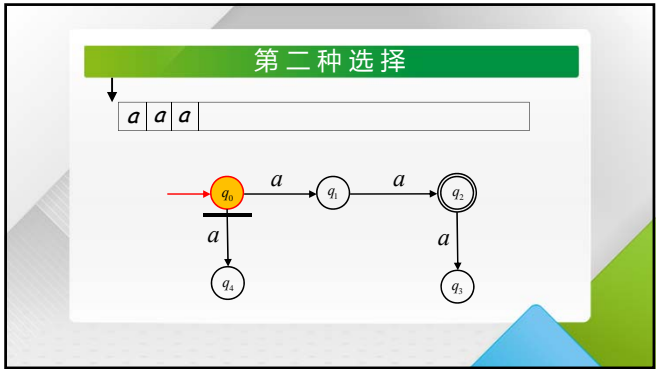
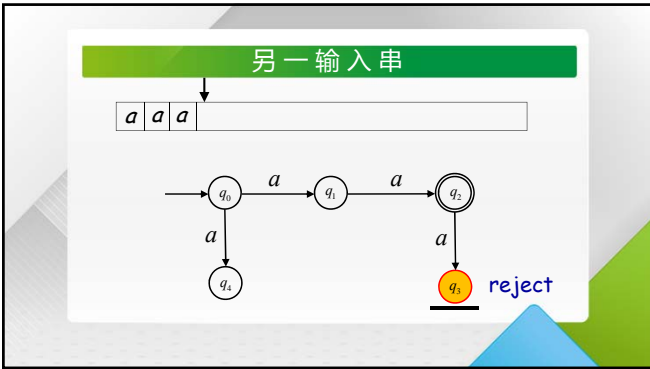
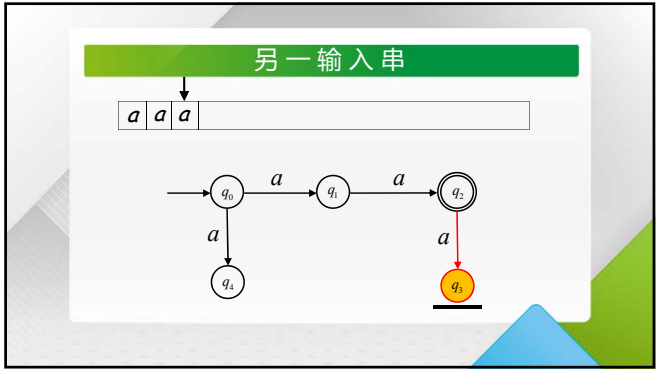
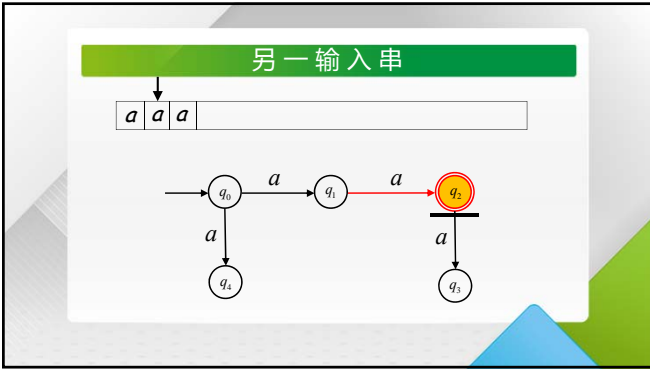
没有迁移

没有迁移





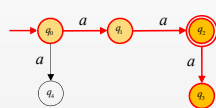




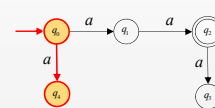
- 字符串 w 被一条路径拒绝：
 - ✓ 字符串输完，但状态不在终态
 - ✓ 字符串无法输完
- 字符串 w 被 NFA 拒绝：
NFA中没有一条路径可以接受字符串 w
- NFA的语言：
NFA接受字符串的集合

输入串 *aaa*

拒绝

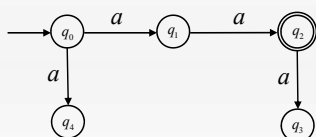


拒绝



NFA所有可能的路径都拒绝字符串 *aaa*
***aaa* 被 NFA 拒绝**

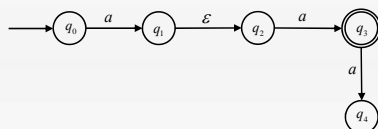
接受的语言: $L=\{aa\}$



非确定有限自动机

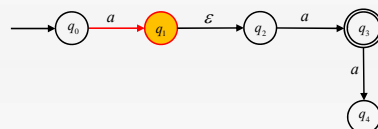
- 非确定有限自动机的概念
- ϵ -转移
- 非确定有限自动机的定义
- 扩展转移函数
- 等价性证明
- 文本搜索

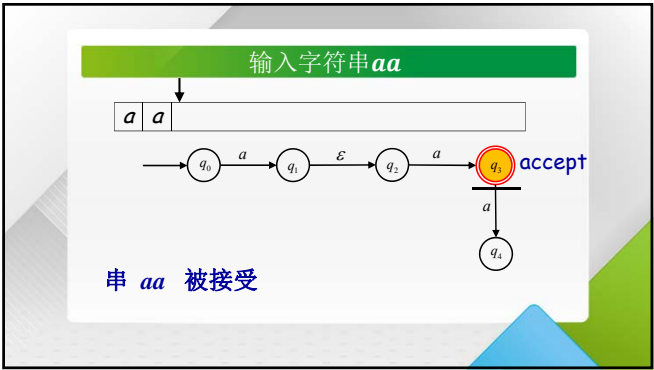
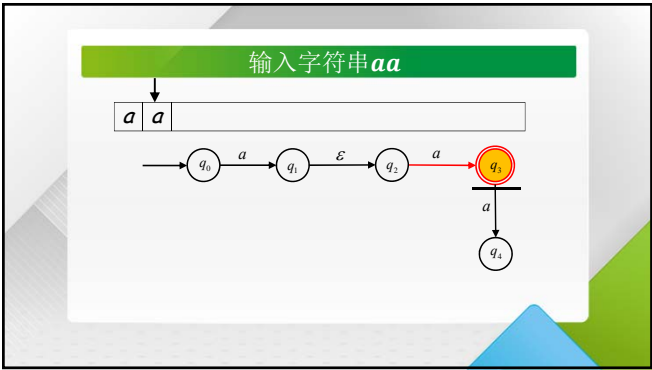
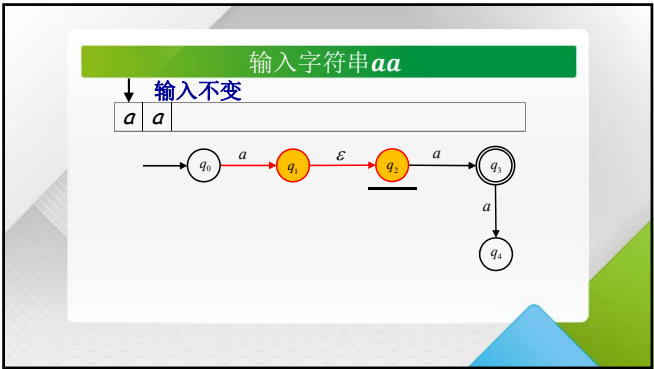
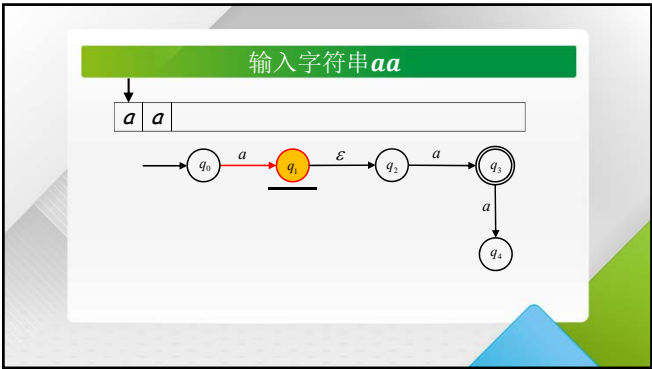
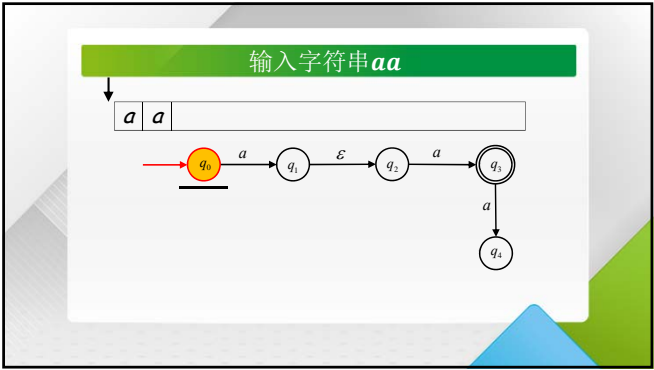
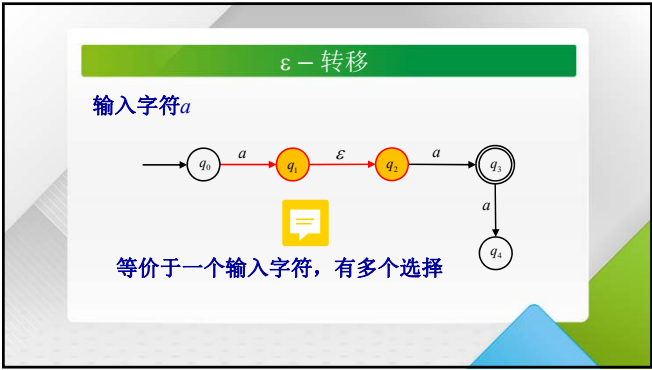
ε - 转移

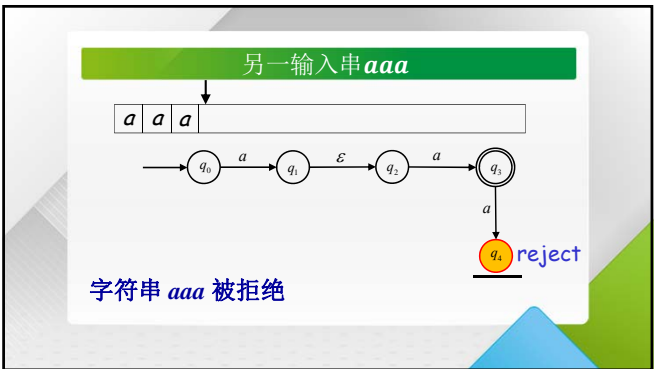
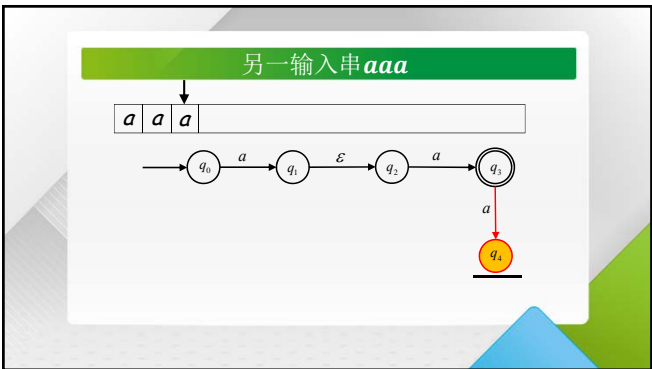
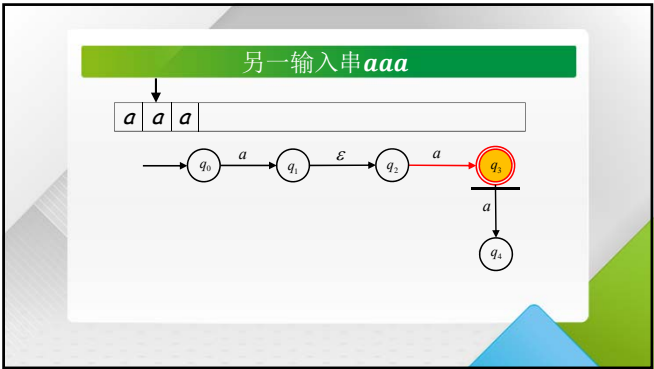
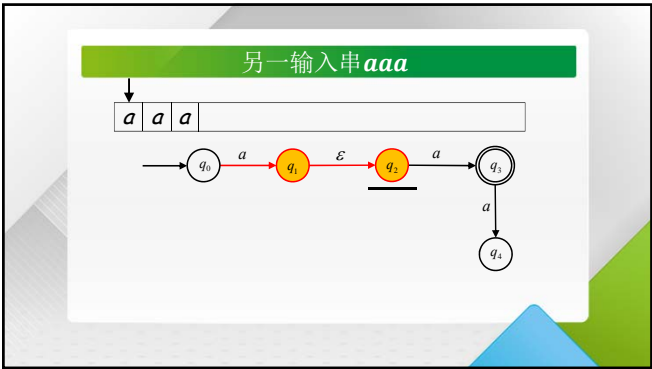
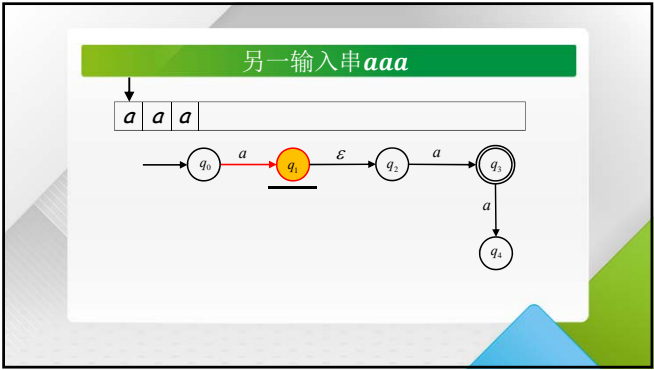
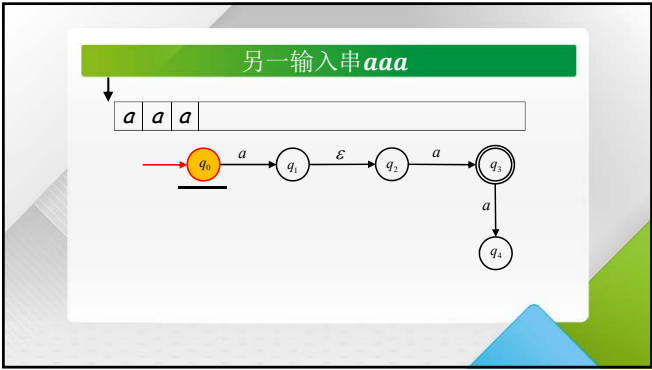


ε - 转移

输入字符 a



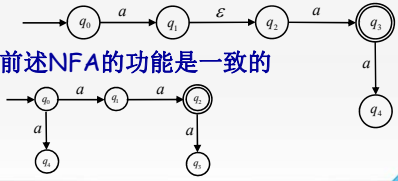




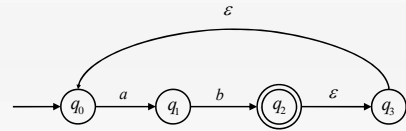
另一输入串 **aaa**

接受的语言: $L=\{aa\}$

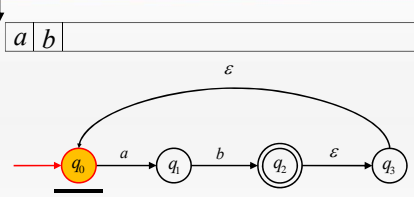
与前述NFA的功能是一致的



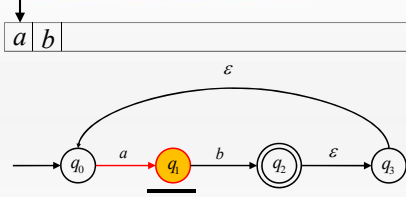
另一个例子



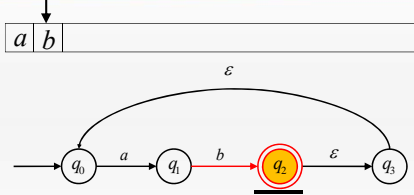
另一个例子



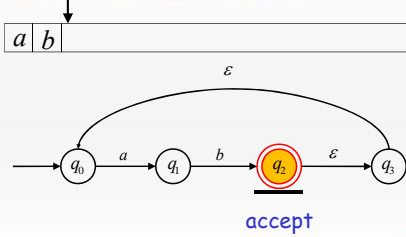
另一个例子

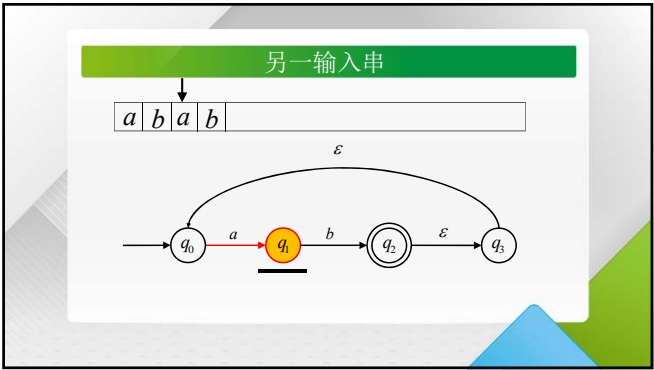
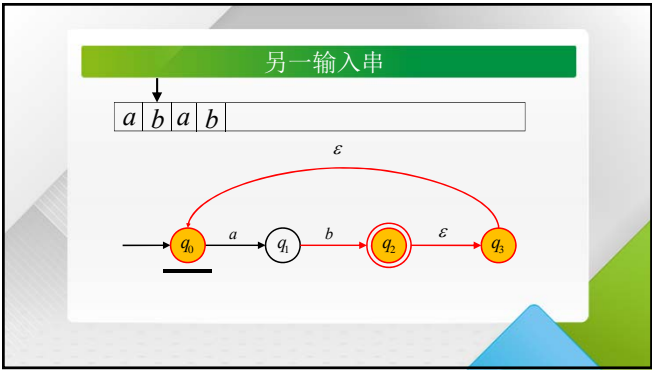
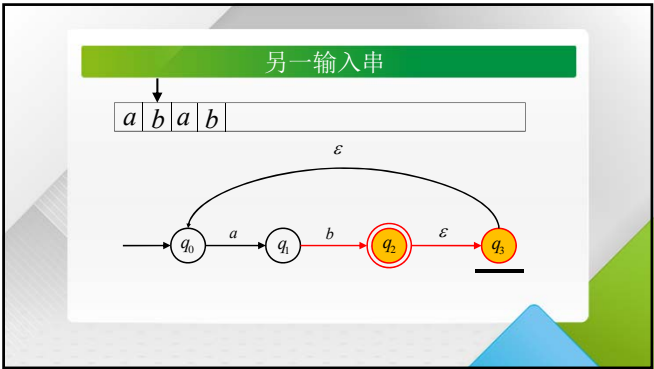
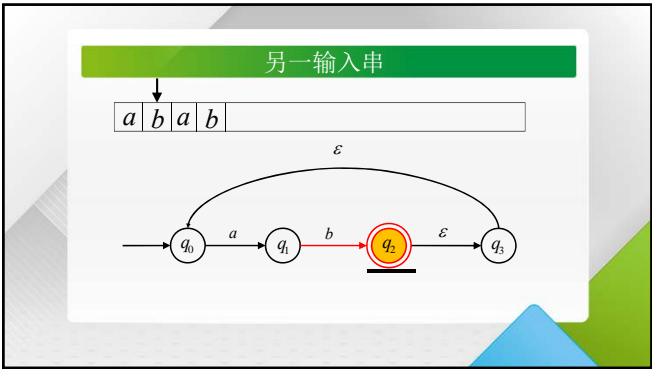
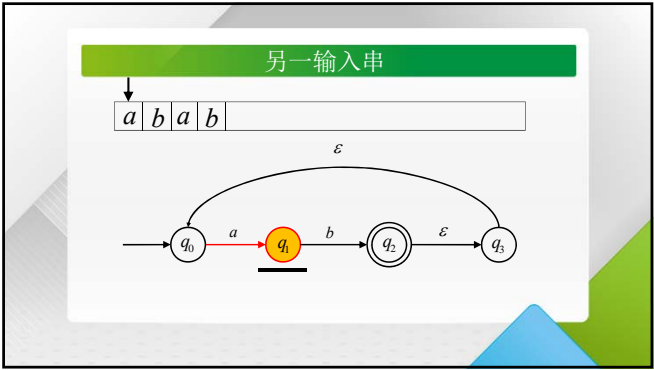
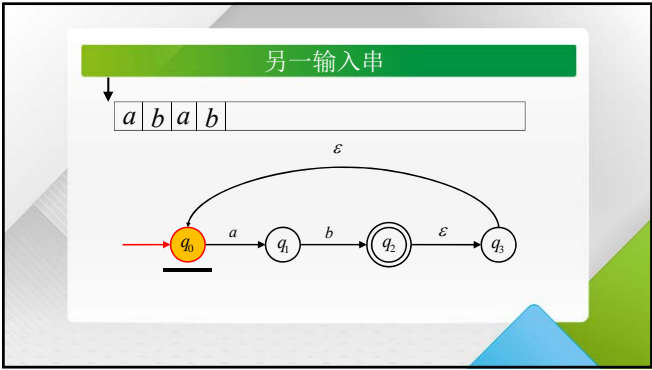


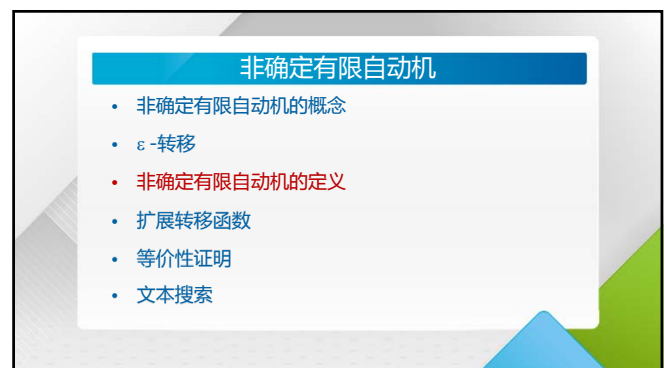
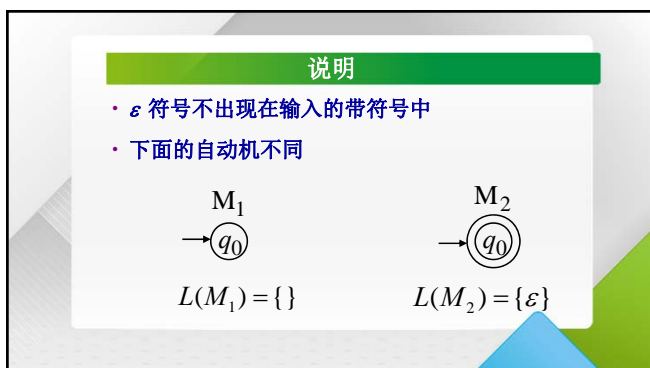
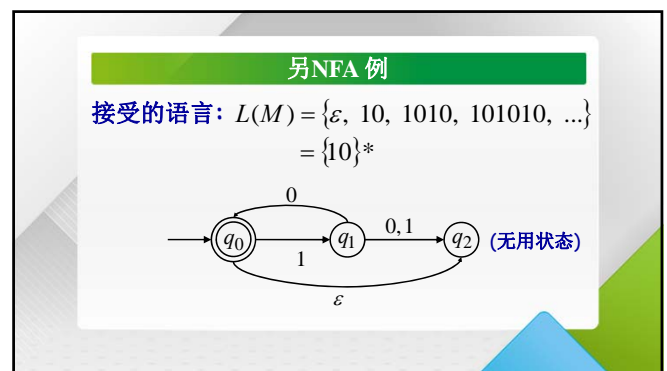
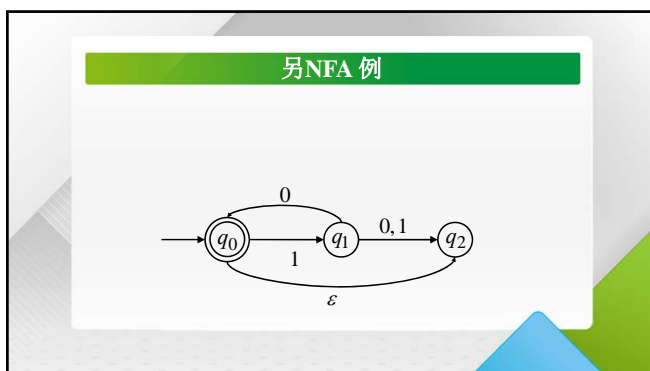
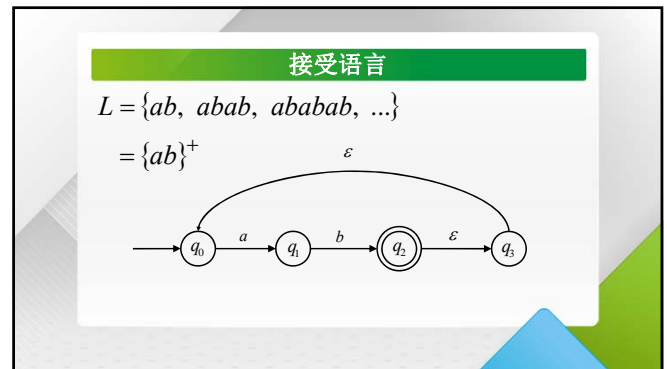
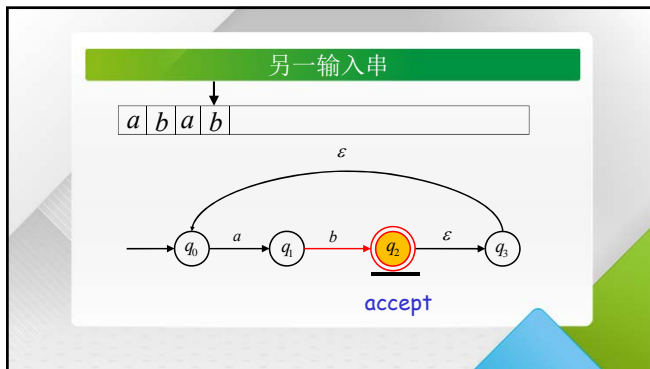
另一个例子



另一个例子







NFA 的定义

非确定有限自动机是五元组 $A = (Q, \Sigma \cup \{\varepsilon\}, \delta, q_0, F)$

有限状态集

输入符号集

转移函数

开始状态

终态集合

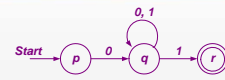
$q_0 \in Q$

$F \subseteq Q$

与 DFA 有三处不同 $\delta: Q \times \Sigma \cup \{\varepsilon\} \rightarrow 2^Q$

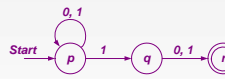
转移图和转移表

(1)



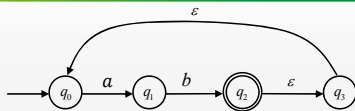
	0	1
$\rightarrow p$	$\{q\}$	\emptyset
q	$\{q\}$	$\{q, r\}$
$*r$	\emptyset	\emptyset

(2)



	0	1
$\rightarrow p$	$\{p\}$	$\{p, q\}$
q	$\{r\}$	$\{r\}$
$*r$	\emptyset	\emptyset

转移图和转移表

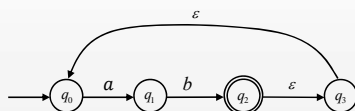


	a	b	ε
$\rightarrow q_0$	$\{q_1\}$	\emptyset	\emptyset
q_1	\emptyset	$\{q_2\}$	\emptyset
q_2	\emptyset	\emptyset	$\{q_3\}$
q_3	\emptyset	\emptyset	$\{q_0\}$

ε - 闭包

- 状态 q 的 ε -闭包是 q (包括 q 自身) 的 ε 路径到达的所有状态, 记为 $EC(q)$ 。
- 归纳定义: 设 $NFA A = (Q, \Sigma, \delta, q_0, F)$, $q \in Q$, $EC(q)$ 满足如下条件:
 - $q \in EC(q)$
 - 若 $p \in EC(q)$ 且 $r \in \delta(p, \varepsilon)$, 则 $r \in EC(q)$
- 集合 S 的 ε -闭包 $EC(S)$ 定义为: $EC(S) = \bigcup_{q \in S} EC(q)$

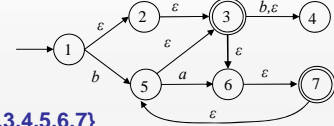
ε - 闭包



$EC(q_2) = \{q_2, q_3, q_0\}$

$EC(q_3) = \{q_3, q_0\}$

ε - 闭包



$EC(1) = \{1, 2, 3, 4, 5, 6, 7\}$

$EC(2) = \{2, 3, 4, 5, 6, 7\}$

$EC(6) = \{5, 6, 7\}$

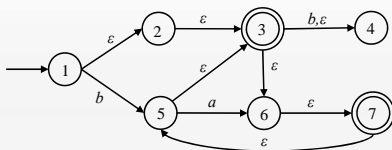
非确定有限自动机

- 非确定有限自动机的概念
- ϵ -转移
- 非确定有限自动机的定义
- 扩展转移函数
- 等价性证明
- 文本搜索

扩展转移函数 δ^*

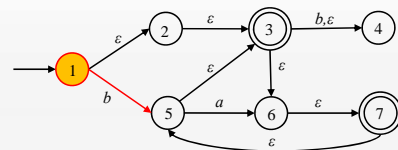
- **NFA状态转移函数**
 $\delta: Q \times \Sigma \cup \{\epsilon\} \rightarrow 2^Q$
- δ 扩展之一:
 $\delta^*: Q \times \Sigma^* \cup \{\epsilon\} \rightarrow 2^Q$
- δ 扩展之二:
 状态转移的 ϵ -闭包

扩展转移函数 δ^*



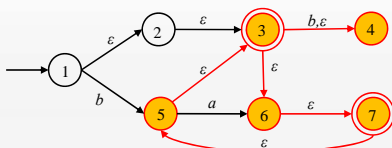
$$\delta^*(1, b)$$

扩展转移函数 δ^*



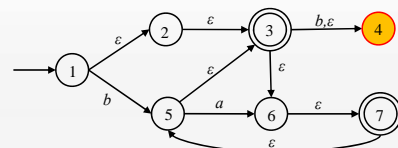
$$\delta^*(1, b)$$

扩展转移函数 δ^*



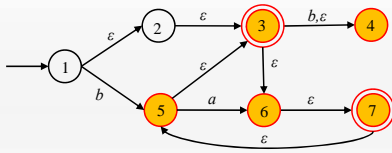
$$\delta^*(1, b) = \{3, 4, 5, 6, 7\}$$

扩展转移函数 δ^*



$$\delta^*(1, bb) = \{4\}$$

扩展转移函数 δ^*



$$\delta^*(1, ba) = \{3, 4, 5, 6, 7\}$$

δ^* 的定义

$q_{j_k} \in \delta^*(q_i, w)$, 则存在一条从 q_i 到 q_{j_k} 的路径 w , 且

$$q_{j_k} \in EC\{q_j\}$$



$$w = \sigma_1 \sigma_2 \cdots \sigma_k$$



NFA语言的形式化定义

NFA M

$$M = (Q, \Sigma, \delta, q_0, F)$$

接受的语言是:

$$L(M) = \{w \in \Sigma^* \mid \delta^*(q_0, w) \cap F \neq \emptyset\}$$

NFA语言的形式化定义

NFA M 接受的语言:

$$L(M) = \{w_1, w_2, w_3, \dots\}$$

满足

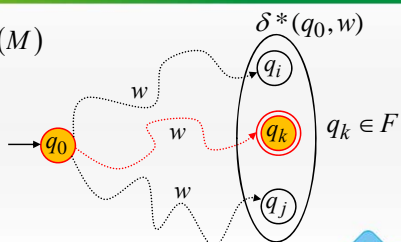
$$\delta^*(q_0, w_m) = \{q_i, q_j, \dots, q_k, \dots\}$$

$$\downarrow$$

$$q_k \in F \text{ (终态)}$$

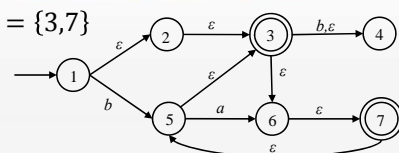
NFA语言的形式化描述

$$w \in L(M)$$

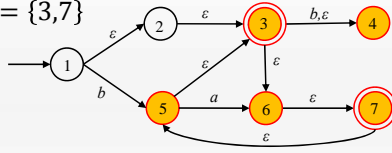


NFA的语言

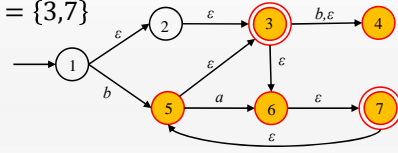
$$F = \{3, 7\}$$



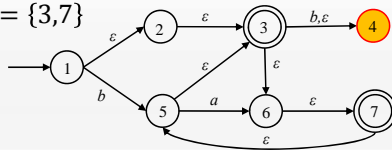
NFA的语言

 $F = \{3,7\}$  $\delta^*(1, b) = \{3,4,5,6,7\}$ $\in F$ $b \in L(M)$

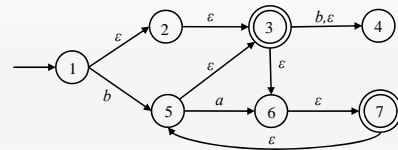
NFA的语言

 $F = \{3,7\}$  $\delta^*(1, baa) = \{3,4,5,6,7\}$ $\in F$ $baa \in L(M)$

NFA的语言

 $F = \{3,7\}$  $\delta^*(1, bb) = \{4\}$ $\notin F$ $bb \notin L(M)$

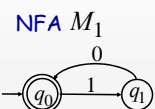
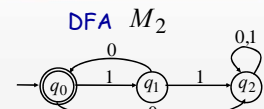
NFA的语言

 $L(M) = b\{a\}^*$

非确定有限自动机

- 非确定有限自动机的概念
- ϵ -转移
- 非确定有限自动机的定义
- 扩展转移函数
- 等价性证明
- 文本搜索

例子

 $L(M_1) = \{10\}^*$  $L(M_2) = \{10\}^*$

自动机的等价性

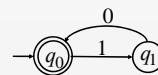
定义:

自动机 M_1 和自动机 M_2 称为等价, 如果两自动机接受的语言相同, 即

$$L(M_1) = L(M_2)$$

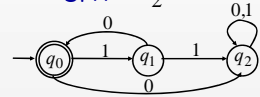
例子

NFA M_1



$$L(M_1) = \{10\}^*$$

DFA M_2



$$L(M_2) = \{10\}^*$$

自动机 M_1 和自动机 M_2 等价

结论

$$\left\{ \begin{array}{l} \text{NFA} \\ \text{接受的语言} \end{array} \right\} = \left\{ \begin{array}{l} \text{正则语言} \\ \text{DFA所接受的语言} \end{array} \right\}$$

DFA所接受的语言

NFA和DFA有着相同的表达能力

需要证明

$$\left\{ \begin{array}{l} \text{NFA} \\ \text{接受的语言} \end{array} \right\} \supseteq \left\{ \begin{array}{l} \text{正则语言} \end{array} \right\}$$

证明:

因为DFA都是NFA, 故结论显然成立

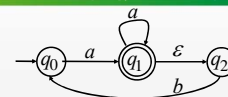
需要证明

$$\left\{ \begin{array}{l} \text{NFA} \\ \text{接受的语言} \end{array} \right\} \subseteq \left\{ \begin{array}{l} \text{正则语言} \end{array} \right\}$$

要证明每一个NFA都可以转化为等价的DFA, 这是一个构造过程

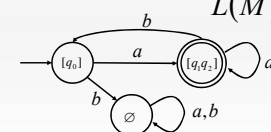
NFA 转换 DFA 例

NFA M



$$L(M) = L(M')$$

DFA M'



NFA 转换 DFA 的算法

给定NFA $M = (Q, \Sigma, \delta, q_0, F)$, 转化为一个等价的DFA M' , 满足

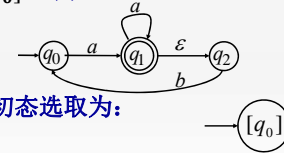
$$L(M) = L(M')$$

DFA M' 状态集合的选取: 设 $Q = \{q_0, q_1, q_2, \dots\}$
则DFA M' 的状态集可以为 Q 的幂集

$$\emptyset, [q_0], [q_1], [q_1, q_2], \dots$$

NFA 转换 DFA 的算法

1. 若NFA的初始状态是 q_0 , 则DFA的初始状态是 $[q_0]$ 。例NFA M :



DFA M' 的初态选取为:

NFA 转换 DFA 的算法

2. 对DFA中的状态 $[q_i, q_j, \dots, q_m]$, 如果在NFA中存在如下迁移:

$$\left. \begin{array}{l} \delta^*(q_i, a), \\ \delta^*(q_j, a), \\ \dots \end{array} \right\} = \{q'_1, q'_j, \dots, q'_m\}$$

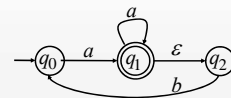
则在DFA中添加迁移:

$$\delta([q_i, q_j, \dots, q_m], a) = [q'_1, q'_j, \dots, q'_m]$$

例

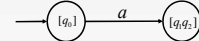
NFA M

$$\delta^*(q_0, a) = \{q_1, q_2\}$$



DFA M'

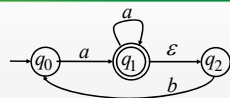
$$\delta([q_0], a) = [q_1, q_2]$$



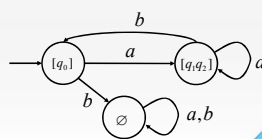
对 Σ 每个字符重复步骤2, 直到无新的转移

例

NFA M



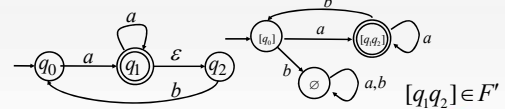
DFA M'



NFA 转换 DFA 的算法

3. DFA接受状态

若状态 q_j 是NFA中的终态, 则DFA 中的状态 $[q_i, q_j, \dots, q_m]$ 是DFA中的终态。例如:



$$[q_1, q_2] \in F'$$

定理

对于NFA M ，通过上述构造过程得到的DFA M' 。则, M 和 M' 是等价的，即：

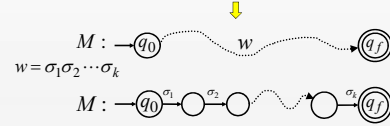
$$L(M) = L(M')$$

证明：需要证明

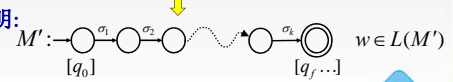
$$L(M) \subseteq L(M') \quad \text{且} \quad L(M) \supseteq L(M')$$

先证明： $L(M) \subseteq L(M')$

对于任意的一个字符串： $w \in L(M)$

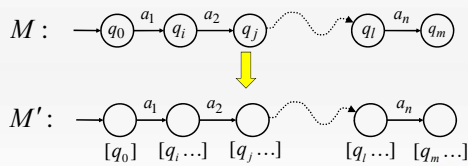


将证明：



一般地，证明在 M 中对任意字符串：

$$v = a_1 a_2 \dots a_n$$

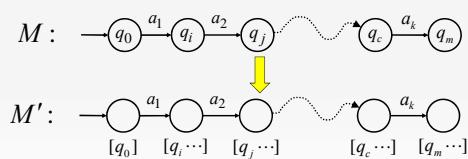


对字符串的长度 $|v|$ 作归纳证明：

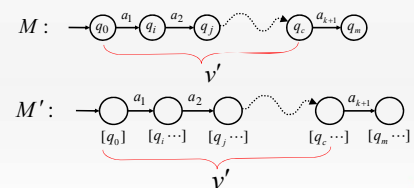
当 $|v|=1$ 时： $v = a_1$



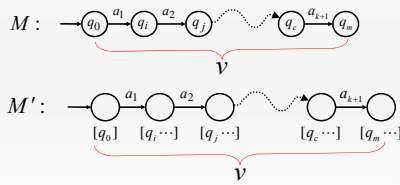
归纳假设： $1 \leq |v| \leq k$ $v = a_1 a_2 \dots a_k$



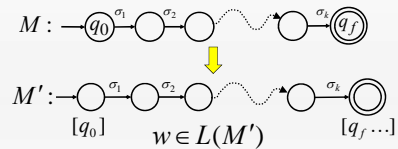
归纳证明： $|v| = k+1$ $v = \underbrace{a_1 a_2 \dots a_k}_{v'} a_{k+1} = v' a_{k+1}$



归纳证明: $|v| = k+1$ $v = \underbrace{a_1 a_2 \cdots a_k}_{v'} a_{k+1} = v' a_{k+1}$



因此, 如果 $w = \sigma_1 \sigma_2 \cdots \sigma_k \in L(M)$



由此证明: $L(M) \subseteq L(M')$
类似可以证明: $L(M) \supseteq L(M')$

证毕

定理说明

- 该定理证明了NFA与DFA接受的语言相同, 都是正则语言。
- 虽然NFA与DFA等价, 从状态的选择可见, 它们表示的复杂度不一样。

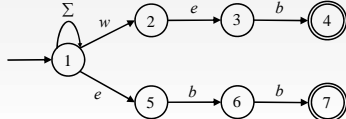
非确定有限自动机

- 非确定有限自动机的概念
- ϵ -转移
- 非确定有限自动机的定义
- 扩展转移函数
- 等价性证明
- 文本搜索

文本搜索

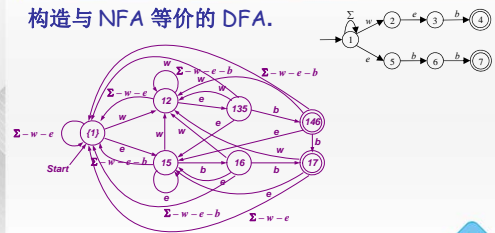
例: 设计一个 NFA 用来在文本中搜索以 *web* 或 *ebb* 为后缀的字符串

下图为一个满足条件的 NFA, 其中 Σ 代表所有 ASCII 字符的集合。



文本搜索

构造与 NFA 等价的 DFA.



课后练习

✧ 必做题:

- P-66 Ex.2.3.2
- P-67 Ex.2.3.3
- P-67 Ex.2.3.4 (b), (c)
- P-72 Ex.2.4.1 (b), (c)
- P-72 Ex.2.4.2
- P-80 Ex.2.5.2
- P-80 Ex.2.5.3 (a)

✧ 思考题:

- P-80 Ex.2.5.3 (b), (c)

2022/3/8

School of Software



109

Thank you