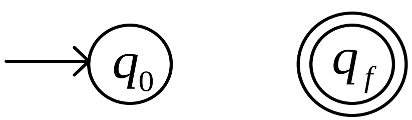
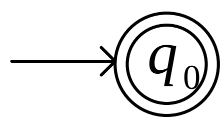
正规文法、正规表达式与有限自动机的等价性

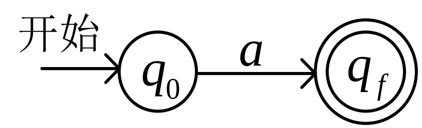
- 前面我们已经证明，从接受语言的范围来说，DFA，NFA，以及带 ε-转换的NFA是等价的。
- 本节我们将分别证明，正规表达式和正规文法同各种FA也是等价的。

2.4.1 正规表达式与有限自动机的等价性

- **定理2.6** 设r为一个正规表达式，那么存在一个带ε-转换的NFA接受L(r)。
- 证明：我们对r中运算符的个数用数学归纳法证明，存在一个带 ε-转换且只有一个终止状态的NFA接受L(r)。
- 1. 归纳基础：当r中运算符个数等于0时，r只可那是下面几种：∅, ε或一个字母 a ，接受它们的NFA分别为：

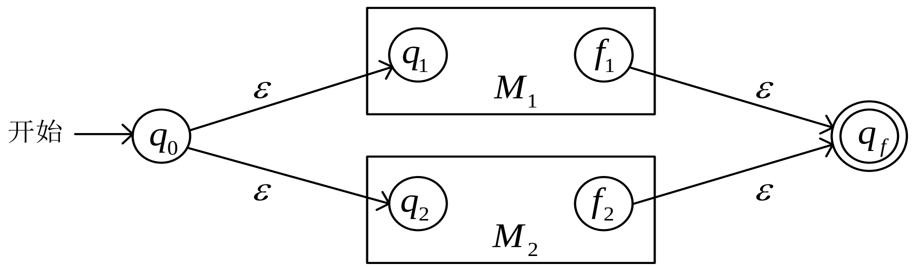




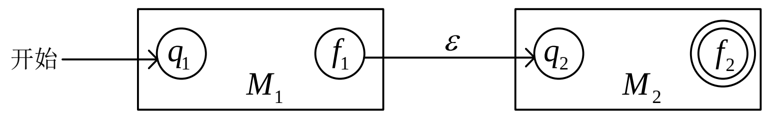


- 1) 接受∅ 2) 接受ε 3)接受a

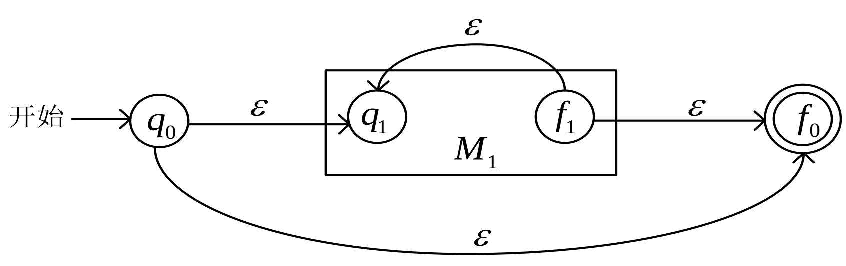
- 归纳步：假设当r中含有不多于k个运算符时结论成立。下面证明当r中含有k+1个运算符时结论也成立。分3种情况讨论。
- 情况1. 设r的最外层运算为“＋”，即$r=r_1+r_2$。此时构造NFA如下：



- 情况2. 设r的最外层运算为连接，即$r=r_1 r_2$。构造NFA如下：



- 情况3. 设r的外层运算为Kleene闭包，即可写成$r=(r_1)^*$的形式。可构造NFA如下：



□

- 例：根据正规式构造等价的NFA。
- 1）$a(a+b)^*b$
- 2）$((a+b)^2)^*$

定理 2.7 设语言 L 被一个 DFA 接受，则 L 也可以被一个正规表达式所表示。

证明:设接受 L 的 DFA 为

$$M=(Q,\Sigma,\delta,q_1,F)$$

其中$Q=\{q_1,q_2,\cdots,q_n\}$。

我们记 $R_{ij}=\{x\in\Sigma^*\mid\delta(q_i,x)=q_j\}$

$$R_{ij}^k=\{x\in R_{ij}\mid\forall y\in \mathrm{Pr}ef(x),y\neq\varepsilon:(\delta(q_1,y)=q_r)\rightarrow(r\leq k)\}$$

其中 Pref(x)表示 x 的前缀，那么易知：

$$R_{ij}^0=\begin{cases}\{a\mid\delta(q_i,a)=q_j\} & \text{当}i\neq j\\ \{a\mid\delta(q_i,a)=q_j\}\cup\{\varepsilon\} & \text{当}i=j\end{cases}$$

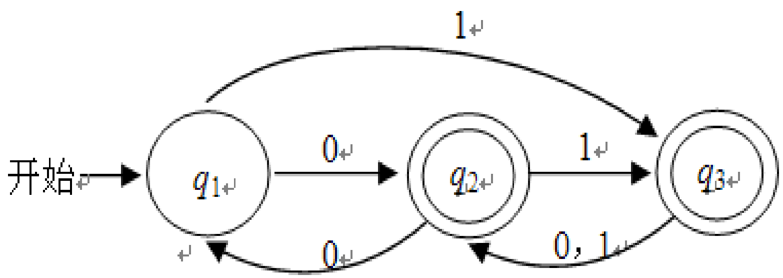
$$R_{ij}^k=R_{ik}^{k-1}(R_{kk}^{k-1})^*R_{kj}^{k-1}\cup R_{ij}^{k-1}$$

用数学归纳法(对 k 进行归纳)可以证明，对 $\forall i, j \in \{1,2,\cdots,n\}$和非负整数 k，都存在正规表达式$r_{ij}^k$表示集合$R_{ij}^k$

由于 $$L(M) = \bigcup_{q_j \in F} R_{1j}^n$$

所以 L(M)可以用一个正规表达式表示。

- 例：构造等价于下列DFA的正规式：



正规文法与有限自动机的等价性

- 正规文法是左线性文法和右线性文法的总称。
- 为了证明正规文法和FA的等价性，我们首先证明每个右线性文法和每个左线性文法所产生的语言都可以被一个带ε-转换的NFA接受，然后证明每个DFA接受的语言都可以由一个右线性文法产生，也可以被一个左线性文法产生。

定理 2.8 设 G 为一个正规文法，那么存在一个带ε-转换的 NFA：M 使得 L(M)=L(G)。

证明：先假设$G=(V,T;P,S)$为一个右线性文法，我们可以构造一个带ε-转换的 NFA：

$$M=(Q,T,\delta,[s],\{[\varepsilon]\})$$

其中：$Q=\{[\alpha]\,|\,\alpha=S$或$\alpha$是G中某产生式右边的一个后缀$\}$

δ定义为：1) 对$A\in V$：$\delta([A],\varepsilon)=\{\alpha|A\rightarrow\alpha$是G的一个产生式$\}$

2) 对$a\in T$且$[a\alpha]\in Q$：$\delta([a\alpha],a)=\{[\alpha]\}$

对推导的步数用数学归纳法可证明

$[\alpha]\in\delta([S],w)$ 当且仅当 $S\overset{*}{\underset{G}{\Rightarrow}}w\alpha$

特别地，当$\alpha=\varepsilon$时，就有：

$$[\varepsilon]\in\delta([S],w) \quad 当且仅当 \quad S\overset{*}{\underset{G}{\Rightarrow}}w$$

即 $$w\in L(M) \quad 当且仅当 \quad w\in L(G)$$

当$G=(V,T;P,S)$是一个左线性文法时，我们可以构造一个右线性文法$G'=(V,T;P',S)$，其中：

$$P'=\{A\to\alpha \mid A\to\alpha^R 是G的一个产生式\}$$

那么易知，G'是一个右线性文法，而且$L(G')=L(G)^R$。对于G'，我们在前面已构造出一个带ε-转换的 NFA：M 接受$L(G')$。注意上面 M 中只有一个终止状态，我们把 M 的终止状态作为初始状态，把 M 的初始状态作为终止状态，把 M 的图形表示中的各条有向边的方向逆转（边上旁标的字符不变），就得到一个新的带ε-转换的 NFA: M'，显然$L(M')=L(M)^R=L(G)$

正规文法→NFA

- 例： 给出正则文法G如下：

  S→0B　B→0B

  B→1S　B→0

试构造等价的DFA。

正规文法→NFA

- 例2.8 设L=0(10)*，易知L可以由下面的右线性文法产生

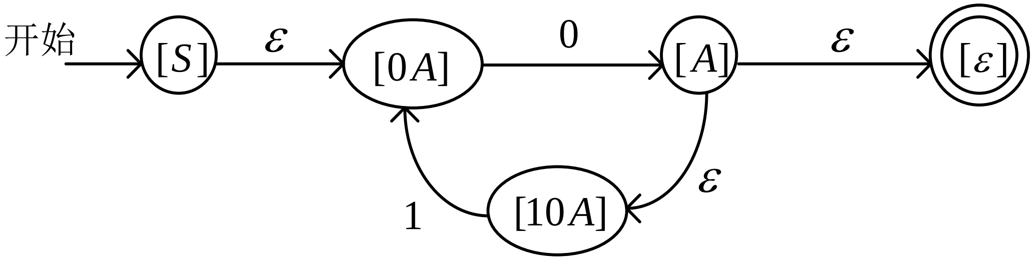
$$G_1: S \to 0A$$

$$A \to 10A \mid \varepsilon$$

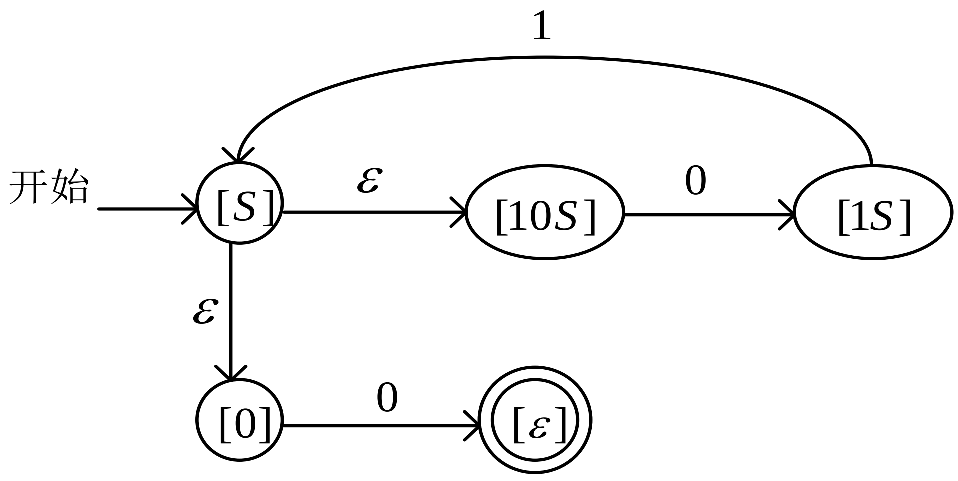
L 也可以由左线性文法G_2产生

$$G_2: S \to S10 \mid 0$$

试分别构造与上述文法等价的 NFA。



根据G_1构造出来的接受 L 的 NFA



根据G_2构造出来的接受 L 的 NFA

NFA→正规文法

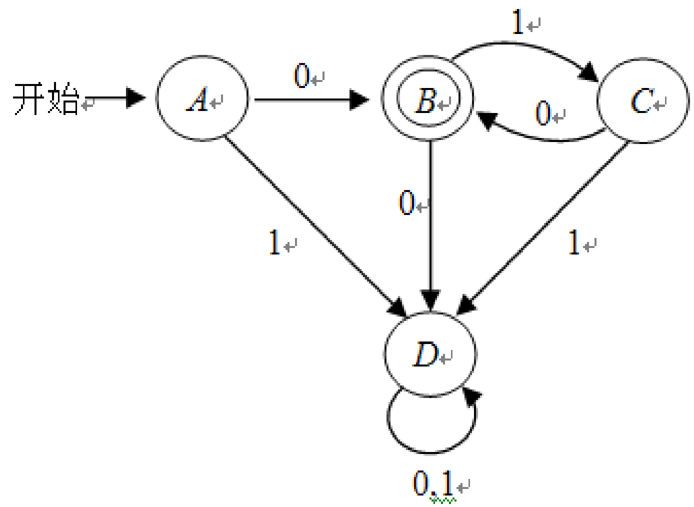
定理 2.9 设$M=(Q,\Sigma,\delta,q_0,F)$为一个 DFA，那么存在一个右线性文法$G_1$使得$L(G_1)=L(M)$，也存在一个左线性文法$G_2$，使得$L(G_2)=L(M)$。

NFA→右线性文法

- 给定NFA：　　$M=(Q, \Sigma, \delta, q_0, F)$
- 构造右线性文法：$G=(Q, \Sigma, P, q_0)$
- 产生式的构造规则：
- 1) $q \rightarrow ap$　如果$(q,a)\rightarrow p$
- 2) $q \rightarrow a$　如果$(q,a)\rightarrow p$且$p\in F$
- 3) $q_0 \rightarrow \varepsilon$　如果 q_0 是终结状态。

NFA→右线性文法

- 例： 给定NFA如下，构造等价的右线性文法。



问 题

- NFA→左线性文法？

2.4.3 右线性文法→正规式

- 给定正规文法，很难直接看清它生成的语言，可以把先求出等价的正规式。
- 方法：将右线性文法变换为正规式的方程组。
- 1. 把文法中的各个变量看作未知正规表达式，终极符看作已知正规表达式。

- 2. 对某变量A，若以A 为左部的全部产生式为
- $A \to xB \mid yC \mid zD$
- 其中x, y, z是终极符号串，B, C, D是变量，则得到正规表达式方程
- $A = xB + yC + zD$
- 3. 这样，整个右线性文法变换为一个正规式方程组，解此方程组可得等价的正规表达式。

- **例2.9** 对于下列文法，求其等价的正规表达式。

$$P: S \to 1S \mid 0A \mid \varepsilon$$
$$A \to 0S \mid 1A$$

- 解：转化为方程组

$$\begin{cases} S = 1S + 0A + \varepsilon \\ A = 0S + 1A \end{cases}$$

- 解此方程组得唯一解

$$S = (1 + 01^*0)^*$$
$$A = 1^*0(1 + 01^*0)^*$$

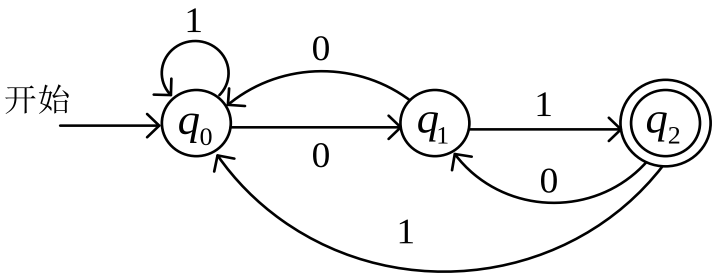
- 因此上述文法所生成的语言是 L((1+01*0)*)

NFA→正规式

- 回顾：去状态法。
- 另一种方法：

NFA→ 右线性文法→正规式

- **例2.10** 求下列DFA的等价正规式。



- 解：首先求出等价的右线性文法

$$q_0 \to 0q_1 \mid 1q_0$$

$$q_1 \to 0q_0 \mid 1q_2 \mid 1$$

$$q_2 \to 0q_1 \mid 1q_0$$

- 然后转化为等价的正规式方程组

$$\begin{cases} q_0 = 0q_1 + 1q_0 \\ q_1 = 0q_0 + 1q_2 + 1 \\ q_2 = 0q_1 + 1q_0 \end{cases}$$

- 解此方程组可得
- q_0=(00+01+1)*01

等价转换方法总结

- 有限自动机之间的等价转换：

  ε-NFA → NFA → DFA

- FA与正规表达式
- FA与右线性文法
- FA与左线性文法
- 右线性文法→正规式方程组→正规式
- 正规式→NFA→右线性文法

研 究

- 左线性文法↔右线性文法

左线性文法↔右线性文法

- 空串对应起始符号，起始符号对应空串。
- $A \to aB \quad \Rightarrow \quad B \to Aa$
- $A \to a \quad \Rightarrow \quad S \to Aa$
- $A \to \varepsilon \quad \Rightarrow \quad S \to A$
- $S \to Aa \quad \Rightarrow \quad A \to a$
- $S \to A \quad \Rightarrow \quad A \to \varepsilon$