

一、判断题（每小题 1 分，共 10 分）

- (1) $\{a, b, c, \dots, z\}$ 是字母表 (✓)
- (2) RL 的同态原像是 RL。 (✓)
- (3) ϕ 不是字母表。 (×)
- (4) Myhill-Nerode 能证明一个语言是正则语言。 (✓)
- (5) 泵引理能证明一个语言是正则语言。 (×)
- (6) $L = \{\varepsilon, 0, 1, 11, 01, 10, 11, 000, \dots\}$, 该语言的句子是所有由 0 和 1 构成的串, 包括空串 ε 。 (✓)
- (7) 设 $L_1, L_2 \subseteq \Sigma^*$, 如果 L_1 是 RL, 则 L_1/L_2 是 RL。 (✓)
- (8) RL 在并、乘积、闭包运算下是封闭的。 (✓)
- (9) $\Sigma = \{0, 1\}, L = \{xwx^T \mid x, w \in \Sigma^+\}$ 是 RL (✓)
- (10) $G = (V, T, P, S)$, 其中 V 为生产式。 (×)

二、简答题（30 分）

1. 简述线性文法、左线性文法、右线性文法。（5 分）

文法 $G = \{V, T, P, S\}$, 如果对于 $\forall \alpha \rightarrow \beta \in P$, 所有 $\alpha \rightarrow \beta$ 均有:

$A \rightarrow w$ 或 $A \rightarrow wBx$ 成立, $A, B \in V, x, w \in T^*$, 则称 G 为线性文法。

文法 $G = \{V, T, P, S\}$, 如果对于 $\forall \alpha \rightarrow \beta \in P$, 所有 $\alpha \rightarrow \beta$ 均有:

$A \rightarrow w$ 或 $A \rightarrow wB$ 成立, 其中 $A, B \in V, w \in T^*$, 则称 G 为右线性文法。

文法 $G = \{V, T, P, S\}$, 如果对于 $\forall \alpha \rightarrow \beta \in P$, 所有 $\alpha \rightarrow \beta$ 均有:

$A \rightarrow w$ 或 $A \rightarrow Bw$ 成立, 其中 $A, B \in V, w \in T^*$, 则称 G 为左线性文法。

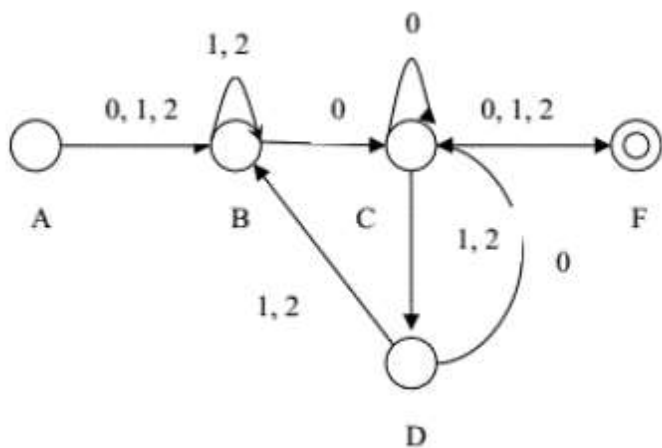
2. 给定如下文法, 请用自然语言描述它们定义的语言。(写出必要步骤) (5 分)

$A \rightarrow 0B \mid 1B \mid 2B$
 $B \rightarrow 0C \mid 1B \mid 2B$
 $C \rightarrow 0E \mid 1D \mid 2D \mid 0 \mid 1 \mid 2$
 $D \rightarrow 0C \mid 1B \mid 2B$
 $E \rightarrow 0E \mid 1D \mid 2D \mid 0 \mid 1 \mid 2$

解: 观察发现 C 和 E 所对应产生式右部是相同的, 所以将文法化简成如下形式:

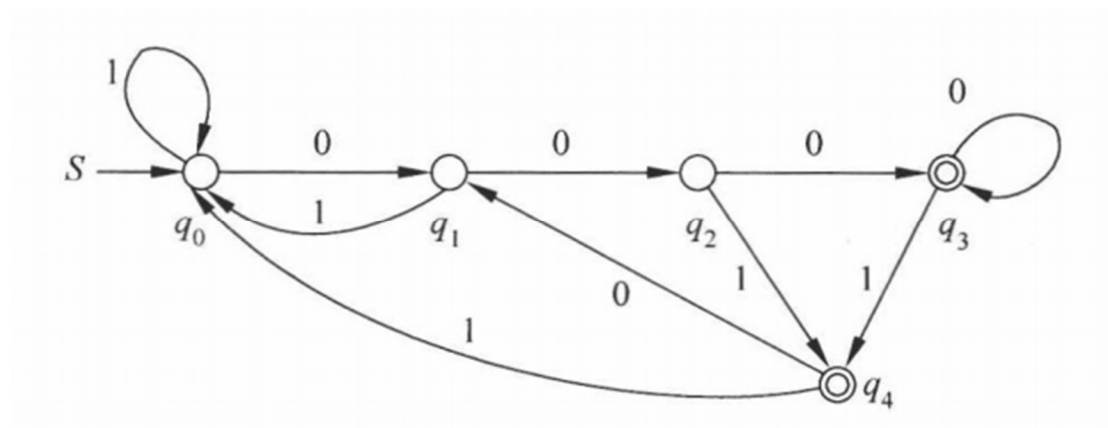
$A \rightarrow 0B \mid 1B \mid 2B$
 $B \rightarrow 0C \mid 1C \mid 2C$
 $C \rightarrow 0D \mid 1D \mid 2D \mid 0 \mid 1 \mid 2$
 $D \rightarrow 0B \mid 1B \mid 2B$

作出状态图如下:

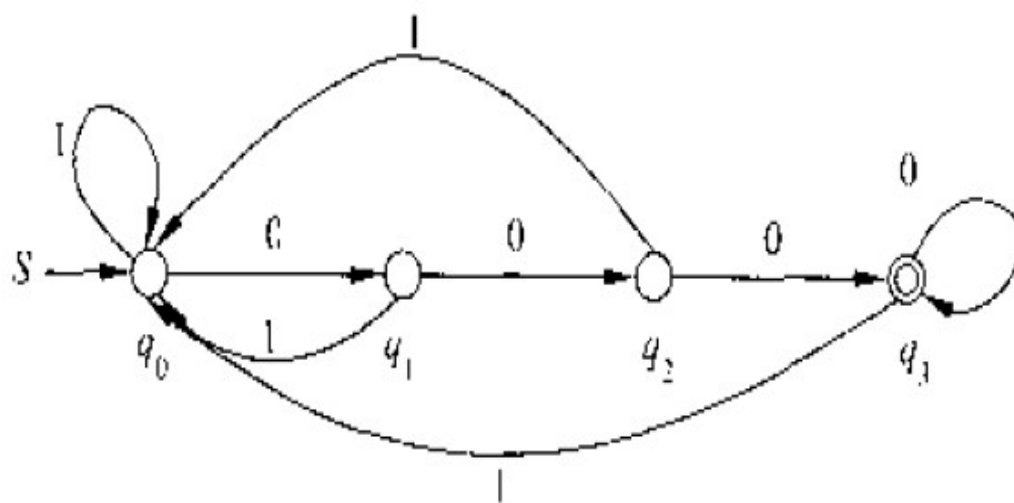


可以看出从初始状态 A 到终态 F，至少要经过 $A \rightarrow B \rightarrow C \rightarrow F$ 的过程，所以字符串的长度至少为 3。而且，到 F 只能经过 C，如果到达 C 后走其它的路径，那么所经过的弧上的字符串都是以 0 为结尾，也就是要回到 C，最后一个字符一定是 0。这样，该文法所确定的语言就是所有倒数第 2 个字符是 0 的串。

3. 画出接受语言 $\{x000|x \in \{0, 1\}^*\} \cup \{x001|x \in \{0, 1\}^*\}$ 的 FA。(5 分)

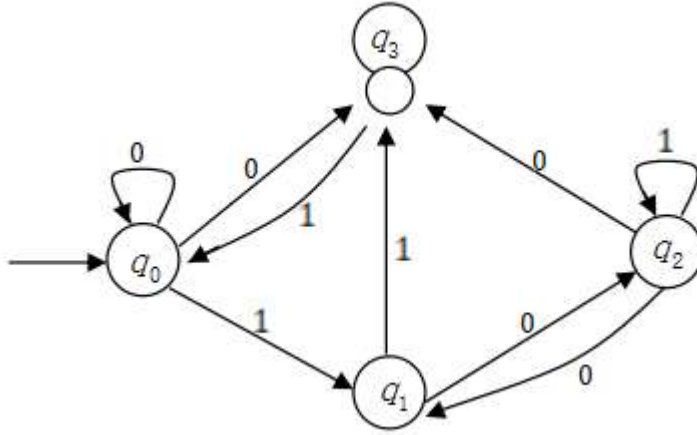


4. (10 分) 对于以下自动机，是否存在能引导 FA 从开始状态到达 q 的字符串的集合 $\text{set}(q)$ 。



$\text{set}(q_0) = \{X \mid X \in \Sigma^*, X = \epsilon \text{ 或者 } X \text{ 以 } 1 \text{ 结尾}\}$
 $\text{set}(q_1) = \{X \mid X \in \Sigma^*, X = 0 \text{ 或者 } X \text{ 以 } 10 \text{ 结尾}\}$
 $\text{set}(q_2) = \{X \mid X \in \Sigma^*, X = 00 \text{ 或者 } X \text{ 以 } 100 \text{ 结尾}\}$
 $\text{set}(q_3) = \{X \mid X \in \Sigma^*, X = 000 \text{ 或者 } X \text{ 以 } 000 \text{ 结尾}\}$

5. 下图是一个 NFA 的状态转移图，给出这个 NFA 的形式定义。（5 分）



$M = (Q, \Sigma, \delta, q_0, F)$ (2 分)

$Q = \{q_0, q_1, q_2, q_3\}, \Sigma = \{0, 1\}$, (3 分)

δ : (5 分)

$\delta(q_0, 0) = \{q_0, q_3\} \quad \delta(q_0, 1) = \{q_1\}$

$\delta(q_1, 0) = \{q_0\} \quad \delta(q_1, 1) = \{q_3\}$

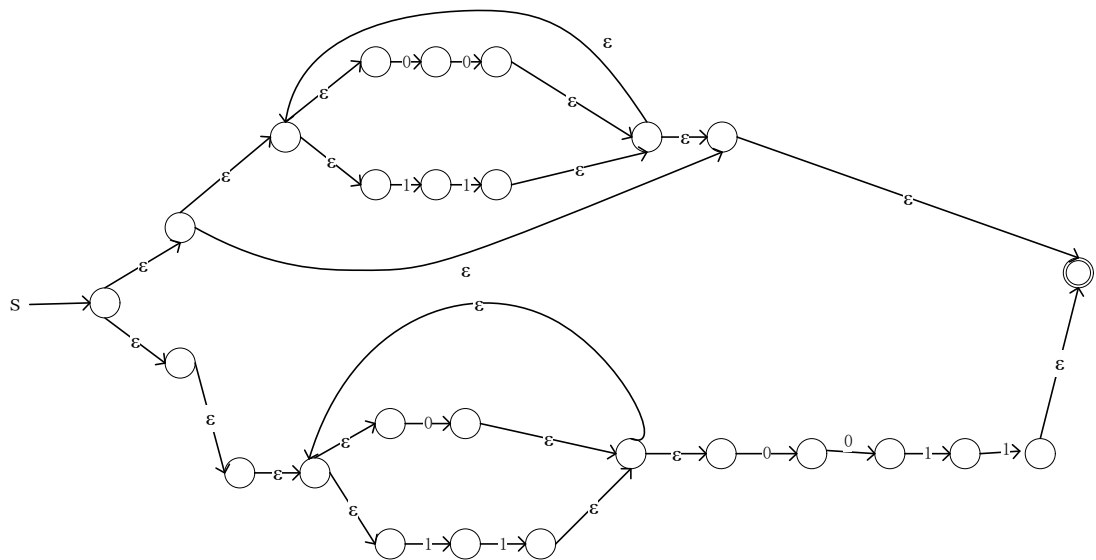
$\delta(q_2, 0) = \{q_1, q_3\} \quad \delta(q_2, 1) = \{q_2\}$

$\delta(q_3, 1) = \{q_0\}$

三、综合题 (60 分)

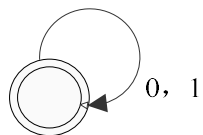
1. 按照书中定理 4-1 证明中所给的方法构造以下正则表达式的等价 FA。(10 分)

$(00+11)^* + (0+11)^+ 0011$

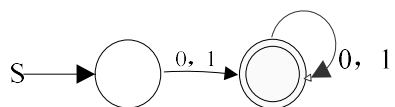


2. 构造下列语言的 DFA 。（10 分）

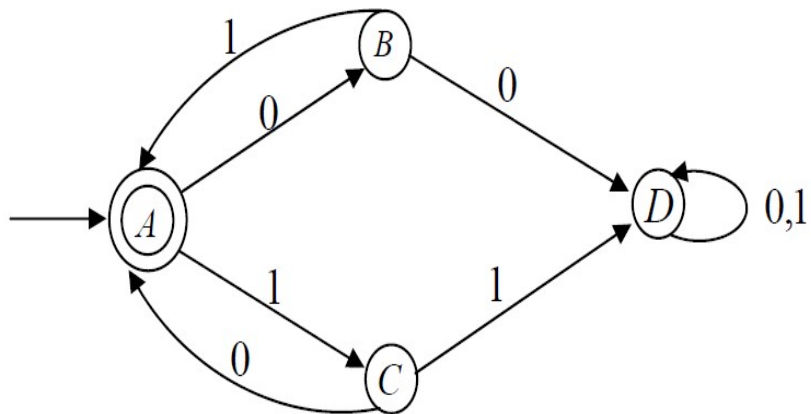
(1) $\{0, 1\}^*$

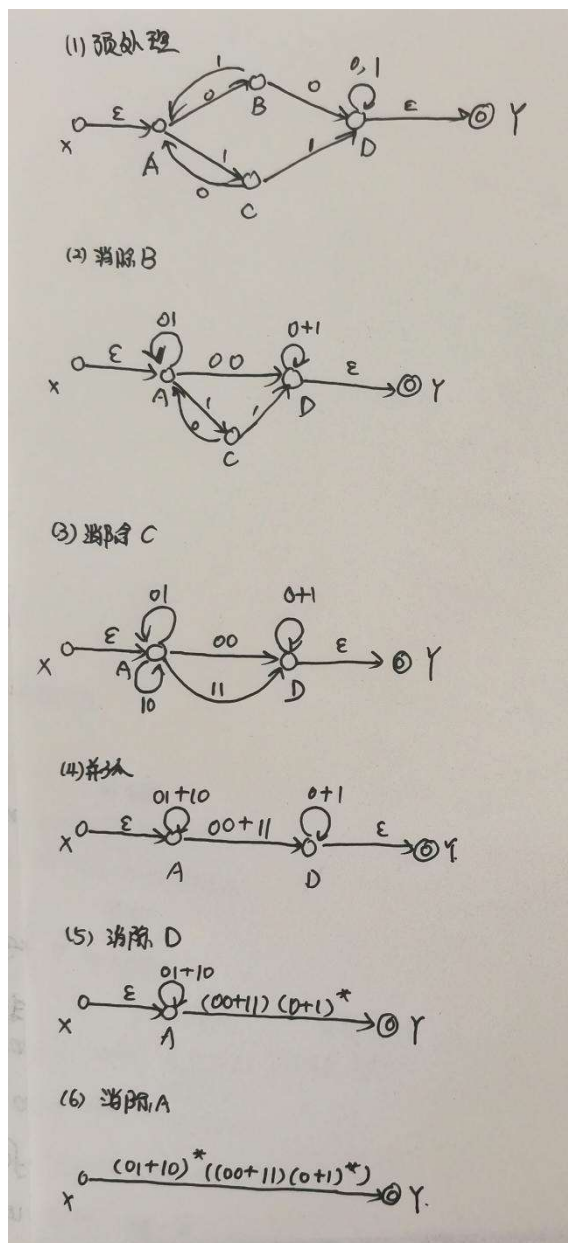


(2) $\{0, 1\}^+$

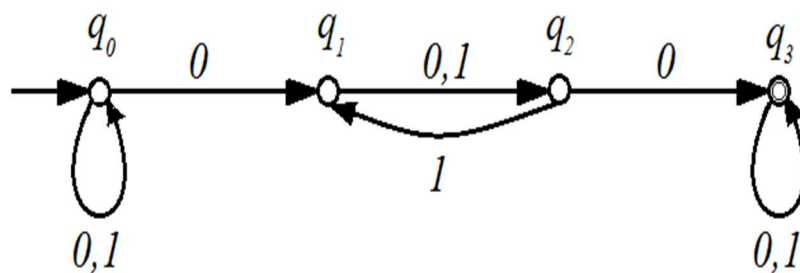


3. 构造等价于下图所示 DFA 的正则表达式。（写出步骤）。（15 分）





4. (10 分) NFA 的状态转移函数图如下所示:



写出这个 NFA 的状态集、输入字符集、状态转移函数、开始状态、终结状态集

$M = (Q, \Sigma, \delta, q_0, q_3)$ (2 分)

$Q = \{q_0, q_1, q_2, q_3\}$, $\Sigma = \{0, 1\}$, (3 分)

δ : (5 分)

$\delta(q_0, 0) = \{q_0, q_1\}$ $\delta(q_0, 1) = \{q_0\}$

$\delta(q_1, 0) = \{q_2\}$ $\delta(q_1, 1) = \{q_2\}$

$\delta(q_2, 0) = \{q_3\}$ $\delta(q_2, 1) = \{q_1\}$

$$\delta(q_3, 0) = \{q_3\} \quad \delta(q_3, 1) = \{q_3\}$$

5. 上下文无关文法化简。(15 分)

$$\begin{aligned} G: & S \rightarrow 0A \mid E \\ & A \rightarrow \varepsilon \mid 0 \mid 0A \mid 1A \mid B \\ & B \rightarrow _C \\ & C \rightarrow 0 \mid 1 \mid 0C \mid 1C \\ & D \rightarrow 1 \mid 1D \mid 2D \mid 3D \mid 4D \mid 5D \mid 6D \\ & E \rightarrow 02E \mid 0E2 \mid 0E3 \mid 0E4 \end{aligned}$$

- (1) 去除文法 G 中的无用符号。(5 分)
- (2) 将(1)得到的文法去除 ε -产生式。(5 分)
- (3) 将(2)得到的文法去除单一产生式。(5 分)

$$\begin{aligned} S &\rightarrow 0A \\ A &\rightarrow \varepsilon \mid 0 \mid 0A \mid 1A \mid B \\ B &\rightarrow _C \\ C &\rightarrow 0 \mid 1 \mid 0C \mid 1C \end{aligned}$$

(2) (5 分)

$$\begin{aligned} S &\rightarrow 0 \mid 0A \\ A &\rightarrow 0 \mid 1 \mid 0A \mid 1A \mid B \\ B &\rightarrow _C \\ C &\rightarrow 0 \mid 1 \mid 0C \mid 1C \end{aligned}$$

(3) (5 分)

$$\begin{aligned} S &\rightarrow 0 \mid 0A \\ A &\rightarrow 0 \mid 1 \mid 0A \mid 1A \mid _C \\ C &\rightarrow 0 \mid 1 \mid 0C \mid 1C \end{aligned}$$

第一章 7、10、22

7. 设 $A = \{1, 2, 3, 4, 5, 6\}$, $B = \{1, 3, 5\}$, $C = \{2, 4, 6\}$ 是在论域 $U = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ 上的集合, 计算下列表达式。

- (1) $A \cap B$ 。
- (2) $(A \cap B) \cup C$ 。
- (3) $(A \cap B) \cup (U - C)$ 。

- (4) $A - B - C$ 。
- (5) $A \times B \times C \times \emptyset$ 。
- (6) $(A \cap B) \cup \overline{A \cup C} \cup \overline{A}$ 。
- (7) $A \times \overline{B} \times \overline{A \cap C}$ 。
- (8) $\overline{A \cup B} \cap (A \cup B) \cup C$ 。

$$1.7 \quad A = \{1,2,3,4,5,6\} \quad B = \{1,3,5\} \quad C = \{2,4,6\} \quad U = \{0,1,2,3,4,5,6,7,8,9\}$$

$$(1). A \cap B$$

$$= \{1,3,5\} = B$$

$$(2). (A \cap B) \cup C$$

$$= \{1,3,5\} \cup \{2,4,6\}$$

$$= \{1,2,3,4,5,6\} = A$$

$$(3). (A \cap B) \cup (U - C)$$

$$= \{1,3,5\} \cup \{0,1,3,5,7,8,9\}$$

$$= \{0,1,3,5,7,8,9\} = \overline{C}$$

$$(4). A - B - C$$

$$= \{2,4,6\} - \{2,4,6\}$$

$$= \Phi$$

$$(5). A \times B \times C \times \Phi$$

$$= \Phi$$

$$\therefore A \times \Phi = \Phi$$

$$(6). (A \cap B) \cup \overline{A \cup C \cup \overline{A}}$$

$$= \{1,3,5\} \cup \{0,7,8,9\} \cup \{0,7,8,9\}$$

$$= \{0,1,3,5,7,8,9\} = \overline{C}$$

$$(7). A \times \overline{B \times \overline{A \cap C}}$$

$$= A \times \overline{B \times \overline{C}}$$

$$= A \times \{(a,b) \mid (a \in B, b \in C) \text{ 或 } (a \in \overline{B}, b \in C) \text{ 或 } (a \in \overline{B}, b \in \overline{C})\}$$

$$= \{(a, b, c) \mid (a \in A, b \in B, c \in C) \text{ 或 } (a \in A, b \in \bar{B}, c \in C) \text{ 或 } (a \in A, b \in \bar{B}, c \in \bar{C})\}$$

$$(8). \overline{A \cup B} \cap (A \cup B) \cup C$$

$$= A \cap A \cup C$$

$$= A \cup C$$

$$= A$$

$$= \{1, 2, 3, 4, 5, 6\}$$

1.10 设 R_1 和 R_2 是集合 $\{a, b, c, d, e\}$ 上的二元关系, 其中,

$$R_1 = \{(a, b), (c, d), (b, d), (b, b), (d, e)\}, R_2 = \{(a, a), (b, c), (d, c), (e, d), (c, a)\}$$

求: $R_1 R_2, R_2 R_1, R_1^+, R_2^+, R_1^*, R_2^*$.

解答:

$$R_1 R_2 = \{(a, c), (c, c), (b, c), (d, d)\}$$

$$R_2 R_1 = \{(a, b), (b, d), (d, d), (e, c), (c, b)\}$$

$$R_1^+ = \{(a, b), (c, d), (b, d), (b, b), (d, e), (a, d), (a, e), (b, e), (c, e)\}$$

$$R_2^+ = \{(a, a), (b, c), (d, c), (c, d), (c, a), (b, a), (d, a), (e, c), (c, a)\}$$

$$R_1^* = R_1^+ \cup \{(a, a), (b, b), (c, c), (d, d), (e, e)\}$$

$$R_2^* = R_2^+ \cup \{(a, a), (b, b), (c, c), (d, d), (e, e)\}$$

22. 设 $\Sigma = \{a, b\}$, 求字符串 $aaaaabbbba$ 的所有前缀的集合, 后缀的集合, 真前缀的集合, 真后缀的集合。

解: (1) 前缀: $\{\epsilon, a, aa, aaa, aaaa, aaaaa, aaaaaa, aaaaaab, aaaaaabb, aaaaaabb, aaaaaabbb, aaaaaabbbb, aaaaaabbbba\}$

(2) 后缀: $\{\epsilon, a, ba, bba, bbba, bbbba, abbbba, aabbbba, aaabbbba, aaaabbbba, aaaaaabbbba\}$

(3) 真前缀: $\{\epsilon, a, aa, aaa, aaaa, aaaaa, aaaaaa, aaaaaab, aaaaaabb, aaaaaabb, aaaaaabbb, aaaaaabbbb\}$

(4) 真后缀: $\{\epsilon, a, ba, bba, bbba, bbbba, abbbba, aabbbba, aaabbbba, aaaabbbba\}$

第二章 3、4、6、7 (第 2 问)

3. 设文法 G 的产生式集如下, 试给出句子 $id + id * id$ 的两个不同的推导和两个不同的归约。

$$E \rightarrow id \mid c \mid +E \mid -E \mid E+E \mid E-E \mid E * E \mid E/E \mid E \uparrow E \mid \text{Fun}(E)$$

推导:

$$(1) E \Rightarrow E+E \Rightarrow E+E * E \Rightarrow E+E * id \Rightarrow E + id * id \Rightarrow id + id * id$$

$$(2) E \Rightarrow E * E \Rightarrow E * id \Rightarrow E + E * id \Rightarrow E + id * id \Rightarrow id + id * id$$

归约:

$$(1) id + id * id \Leftarrow E + id * id \Leftarrow E + E * id \Leftarrow E + E * E \Leftarrow E + E \Leftarrow E$$

$$(2) id + id * id \Leftarrow E + id * id \Leftarrow E + E * id \Leftarrow E * id \Leftarrow E * E \Leftarrow E$$

4. 设文法 G 的产生式集如下, 试给出句子 $aaabbbccc$ 的至少两个不同的推导和至少两个不同的归约。

$S \rightarrow aBC \mid aSBC$

$aB \rightarrow ab$

$bB \rightarrow bb$

$CB \rightarrow BC$

$bC \rightarrow bc$

$cC \rightarrow cc$

推导:

1) $S \Rightarrow aSBC$	使用产生式 $S \rightarrow aSBC$
$\Rightarrow aaSBCBC$	使用产生式 $S \rightarrow aSBC$
$\Rightarrow aaaBCBCBC$	使用产生式 $S \rightarrow aBC$
$\Rightarrow aaabCBCBC$	使用产生式 $aB \rightarrow ab$
$\Rightarrow aaabBCCBC$	使用产生式 $CB \rightarrow BC$
$\Rightarrow aaabbbCCBC$	使用产生式 $bB \rightarrow bb$
$\Rightarrow aaabbbCBCC$	使用产生式 $CB \rightarrow BC$
$\Rightarrow aaabbbBCCC$	使用产生式 $CB \rightarrow BC$
$\Rightarrow aaabbbbCCC$	使用产生式 $bB \rightarrow bb$
$\Rightarrow aaabbbbccc$	使用产生式 $bC \rightarrow bc$
$\Rightarrow aaabbbbccc$	使用产生式 $cC \rightarrow cc$

2) $S \Rightarrow aSBC$	使用产生式 $S \rightarrow aSBC$
$\Rightarrow aaSBCBC$	使用产生式 $S \rightarrow aSBC$
$\Rightarrow aaaBCBCBC$	使用产生式 $S \rightarrow aBC$
$\Rightarrow aaaBCCBC$	使用产生式 $CB \rightarrow BC$
$\Rightarrow aaaBBCCBC$	使用产生式 $CB \rightarrow BC$
$\Rightarrow aaabbbCBCC$	使用产生式 $aB \rightarrow ab$
$\Rightarrow aaabbbBCCC$	使用产生式 $CB \rightarrow BC$
$\Rightarrow aaabbbbCCC$	使用产生式 $bB \rightarrow bb$
$\Rightarrow aaabbbbccc$	使用产生式 $bC \rightarrow bc$
$\Rightarrow aaabbbbccc$	使用产生式 $cC \rightarrow cc$

归约为推导的逆过程

6. 设文法 G 的产生式集如下, 请给出 G 的每个语法范畴代表的集合。

$S \rightarrow aSa \mid aaSaa \mid aAa$

$A \rightarrow bA \mid bbbA \mid bB$

$B \rightarrow cB \mid cC$

$C \rightarrow ccC \mid DD$

$D \rightarrow dD \mid d$

解:

$$\text{set}(D) = \{d\}^+$$

$$\text{set}(C) = \{c^{2n}d^m \mid m \geq 2, n \geq 0\}$$

$$\text{set}(B) = \{c^n d^m \mid m \geq 2, n \geq 1\}$$

$$\text{set}(A) = \{b^p c^n d^m \mid p \geq 1, m \geq 2, n \geq 1\}$$

$$\text{set}(S) = \{a^q b^p c^n d^m a^q \mid p \geq 1, m \geq 2, n \geq 1, q \geq 1\}$$

7. 给定如下文法, 请用自然语言描述它们定义的语言。

$$(2) A \rightarrow 0B \mid 1B \mid 2B$$

$$B \rightarrow 0C \mid 1C \mid 2C$$

$$C \rightarrow 0D \mid 1D \mid 2D \mid 0 \mid 1 \mid 2$$

$$D \rightarrow 0B \mid 1B \mid 2B$$

解: 该语言的句子是字母表 $\Sigma = \{0, 1, 2\}$ 上所有长度为 3 的倍数的字符串, 且非空。

第三章 2 (1-6)、10 (1-6)、20

2. 构造识别下列语言的 DFA (给出相应 DFA 的形式描述或者画出它们的状态转移图)。

(1) $\{0, 1\}^*$ 。

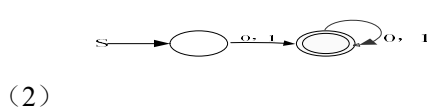
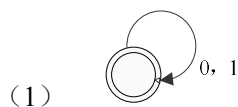
(2) $\{0, 1\}^+$ 。

(3) $\{x \mid x \in \{0, 1\}^+ \text{ 且 } x \text{ 中不含形如 } 00 \text{ 的子串}\}$ 。

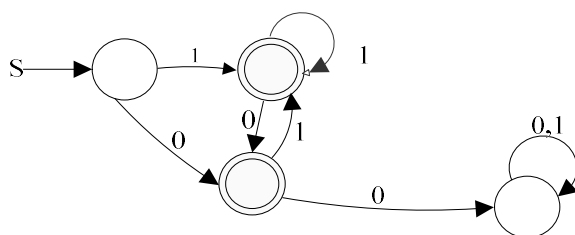
(4) $\{x \mid x \in \{0, 1\}^* \text{ 且 } x \text{ 中不含形如 } 00 \text{ 的子串}\}$ 。

(5) $\{x \mid x \in \{0, 1\}^+ \text{ 且 } x \text{ 中含形如 } 10110 \text{ 的子串}\}$ 。

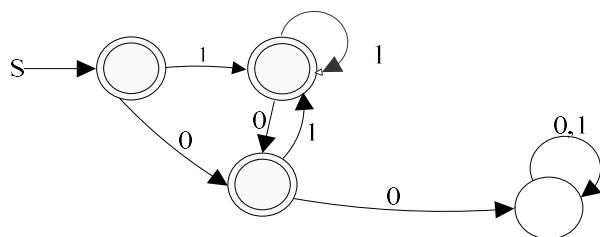
(6) $\{x \mid x \in \{0, 1\}^+ \text{ 且 } x \text{ 中不含形如 } 10110 \text{ 的子串}\}$ 。



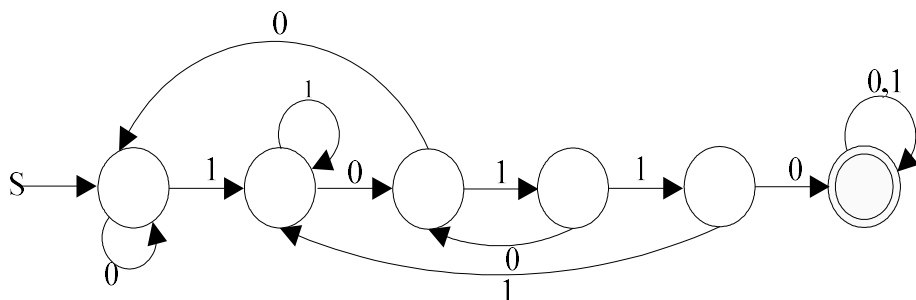
(3) (设置一个陷阱状态, 一旦发现 00 的子串, 就进入陷阱状态)



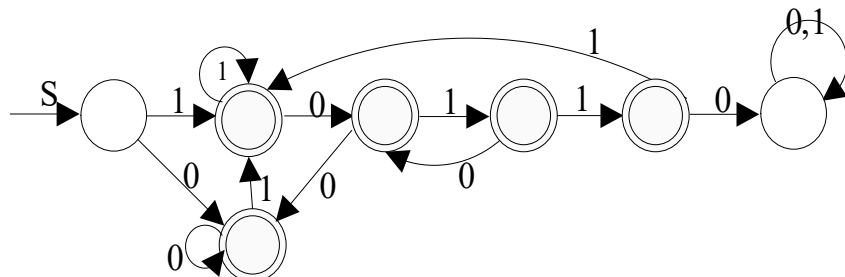
(4)



(5)



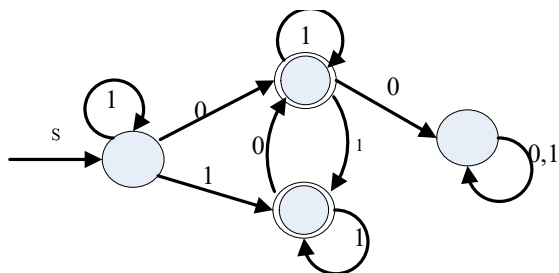
(6) (设置一个陷阱状态，一旦发现有 10110 的子串，就进入陷阱状态)



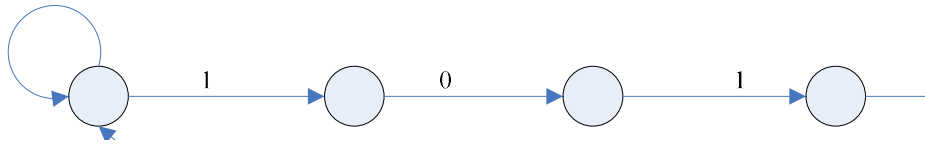
10. 构造识别下列语言的 NFA。

- (1) $\{x | x \in \{0,1\}^+ \text{ 且 } x \text{ 中不含形如 } 00 \text{ 的子串}\}$ 。
- (2) $\{x | x \in \{0,1\}^+ \text{ 且 } x \text{ 中含形如 } 10110 \text{ 的子串}\}$ 。
- (3) $\{x | x \in \{0,1\}^+ \text{ 且 } x \text{ 中不含形如 } 10110 \text{ 的子串}\}$ 。
- (4) $\{x | x \in \{0,1\}^+ \text{ 和 } x \text{ 的倒数第十个字符是 } 1, \text{ 且以 } 01 \text{ 结尾}\}$ 。
- (5) $\{x | x \in \{0,1\}^+ \text{ 且 } x \text{ 以 } 0 \text{ 开头以 } 1 \text{ 结尾}\}$ 。
- (6) $\{x | x \in \{0,1\}^+ \text{ 且 } x \text{ 中至少含两个 } 1\}$ 。

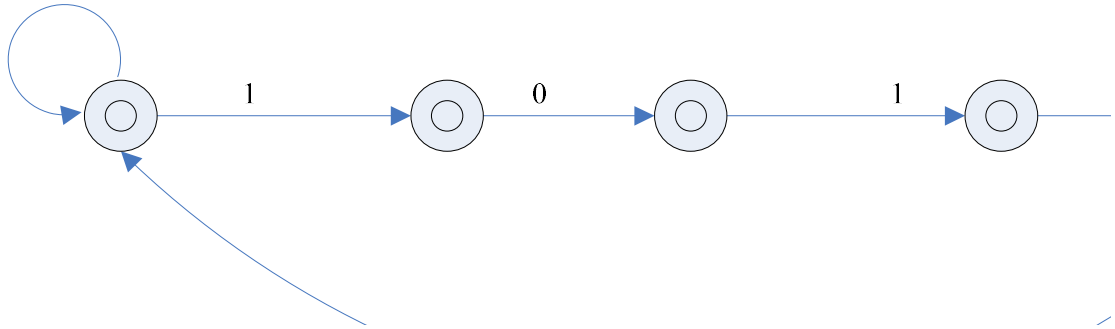
(1)



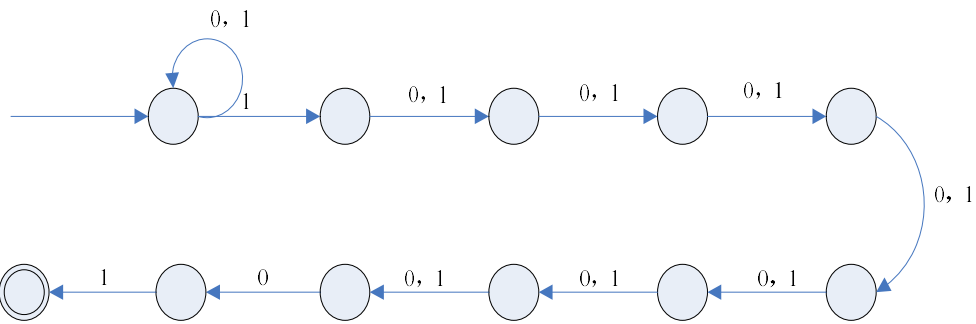
(2)



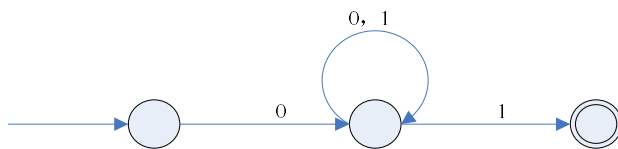
(3)



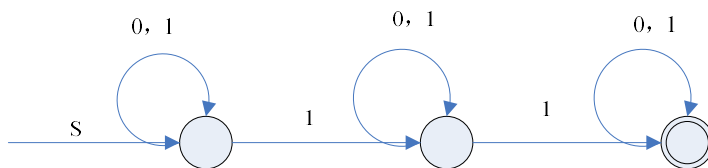
(4)



(5)



(6)



20. 构造图 3-20 所给 DFA 对应的右线性文法。

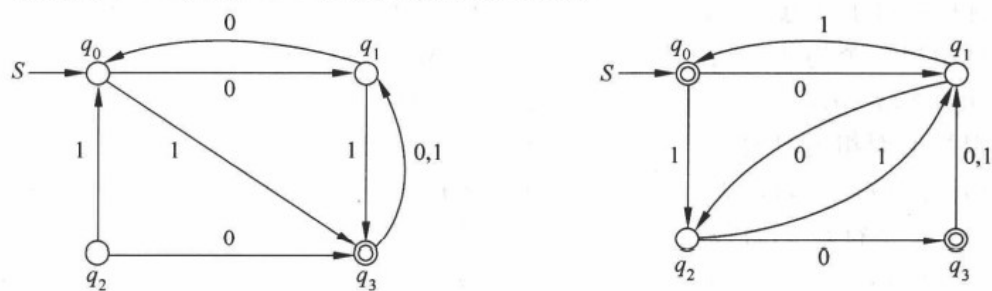


图 3-20 两个不同的 DFA

对图 1:

$q_0 \rightarrow 0q_1 \mid 1q_3 \mid 1$
 $q_1 \rightarrow 0q_0 \mid 1q_3 \mid 1$
 $q_2 \rightarrow 0q_3 \mid 1q_0 \mid 0$
 $q_3 \rightarrow 0q_1 \mid 1q_2$
 设 q_0, q_1, q_2, q_3 分别对应字符 S, A, B, C
 构造文法 $G = (V, T, P, S)$, 其中
 $V = \{S, A, B, C\}, T = \{1, 0\}$
 $P:$
 $S \rightarrow 0A \mid 1C \mid 1$
 $A \rightarrow 0S \mid 1C \mid 1$
 $B \rightarrow 0C \mid 0 \mid 1A$
 $C \rightarrow 0A \mid 1A$

对图 2: 构造文法 $G = (V, T, P, S)$, 其中
 $V = \{S, A, B, C\}, T = \{1, 0\}$

$P:$

$S \rightarrow 0A \mid 1B$
 $A \rightarrow 1S \mid 1 \mid 0B$
 $B \rightarrow 0C \mid 0 \mid 1A$
 $C \rightarrow 0A \mid 1A$

第四章 1 (1-6)、5 (1-4)、6 (图 4-24 图 4-25)

1. 写出表示下列语言的正则表达式。

- (1) $\{0,1\}^*$ 。
- (2) $\{0,1\}^+$ 。
- (3) $\{x \mid x \in \{0,1\}^+ \text{ 且 } x \text{ 中不含形如 } 00 \text{ 的子串}\}$ 。
- (4) $\{x \mid x \in \{0,1\}^* \text{ 且 } x \text{ 中不含形如 } 00 \text{ 的子串}\}$ 。
- (5) $\{x \mid x \in \{0,1\}^+ \text{ 且 } x \text{ 中含形如 } 10110 \text{ 的子串}\}$ 。
- (6) $\{x \mid x \in \{0,1\}^+ \text{ 且 } x \text{ 中不含形如 } 10110 \text{ 的子串}\}$ 。

(1) $\{0,1\}^*$ 。

解：所求正则表达式为： $(0+1)^*$ 。

(2) $\{0,1\}^+$ 。

解：所求正则表达式为： $(0+1)^+$ 。

(3) $\{x \mid x \in \{0,1\}^+ \text{ 且 } x \text{ 中不含形如 } 00 \text{ 的子串}\}$ 。

解：根据第三章构造的 FA，可得所求正则表达式为： $1^*(01^+)^*(01+0+1)$ 。

(4) $\{x \mid x \in \{0,1\}^* \text{ 且 } x \text{ 中不含形如 } 00 \text{ 的子串}\}$ 。

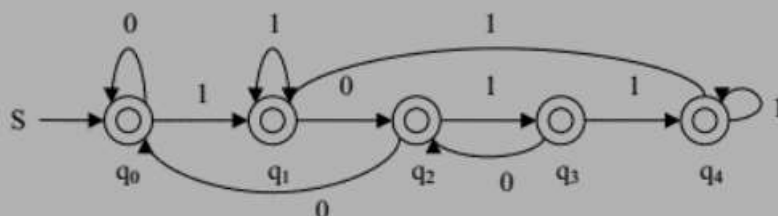
解：根据上题的结果，可得所求正则表达式为： $\epsilon + 1^*(01^+)^*(01+0+1)$ 。

(5) $\{x \mid x \in \{0,1\}^+ \text{ 且 } x \text{ 中含形如 } 10110 \text{ 的子串}\}$ 。

解：所求正则表达式为： $(0+1)^*10110(0+1)^*$ 。

(6) $\{x \mid x \in \{0,1\}^+ \text{ 且 } x \text{ 中不含形如 } 10110 \text{ 的子串}\}$ 。

解：根据第三章的习题，接受 x 的 FA 为：



要求该 FA 对应的正则表达式，分别以 q_0 、 q_1 、 q_2 、 q_3 、 q_4 为终结状态考虑：

q_0 为终态时的正则表达式： $(0^*(11^*0(10)^*(\epsilon + 111^*11^*0(10)^*0)^*)^*)^*$

q_1 为终态时的正则表达式： $0^*1(1^*(0(10)^*111^*1)^*(0(10)^*00^*1)^*)^*$

q_2 为终态时的正则表达式： $0^*11^*0((10)^*(111^*11^*0)^*(00^*11^*0)^*)^*$

q_3 为终态时的正则表达式： $0^*11^*0(10)^*1(11^*11^*0((10)^*(00^*11^*0)^*)^*1)^*$

q_4 为终态时的正则表达式： $0^*11^*0(10)^*11(1^*(11^*0((00^*11^*0)^*(10)^*)^*11)^*)^*$

将以上 5 个正则表达式用 “+” 号相连，就得到所要求的正则表达式。

5. 构造下列正则表达式的等价 FA。

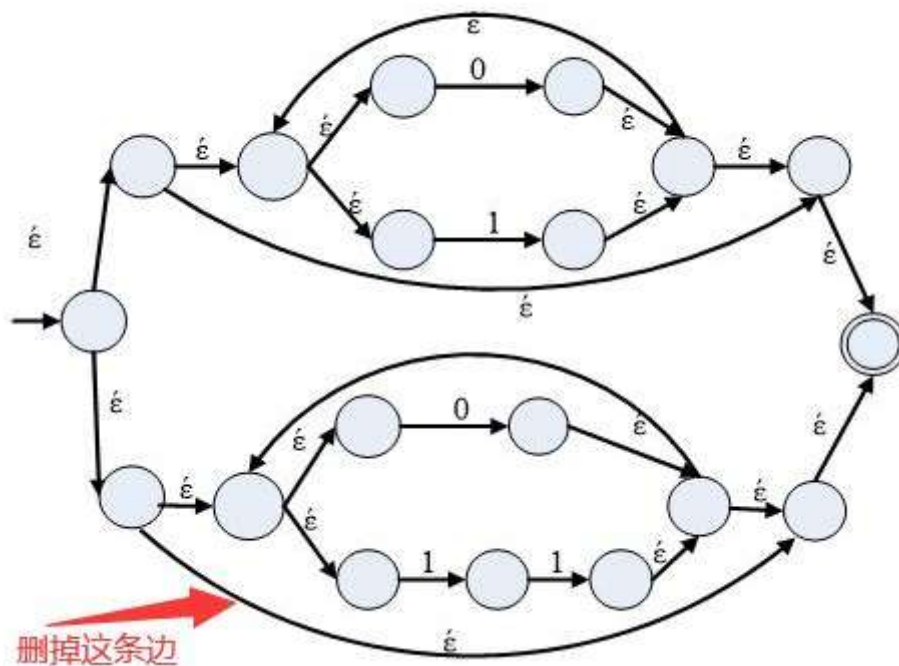
(1) $(0+1)^* + (0+11)^+$ 。

(2) $00(0+1)^*((01)^* + 010)00$ 。

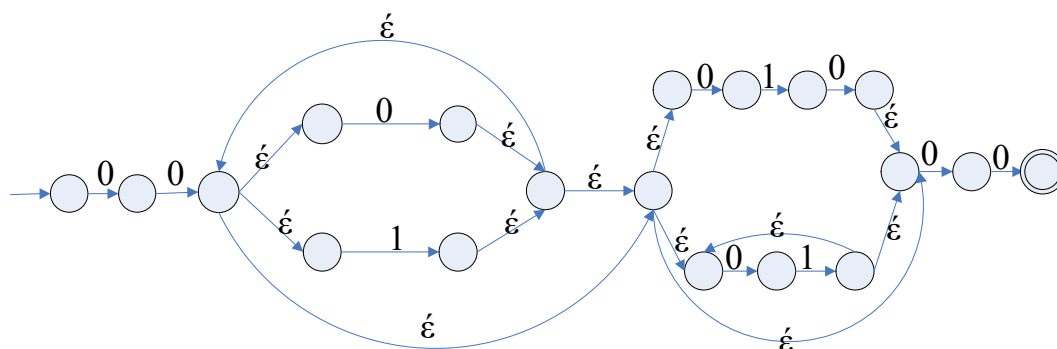
(3) $(1+01+001)^*(\epsilon+0+00)$ 。

(4) $((0+1)(0+1))^* + ((0+1)(0+1)(0+1))^*$ 。

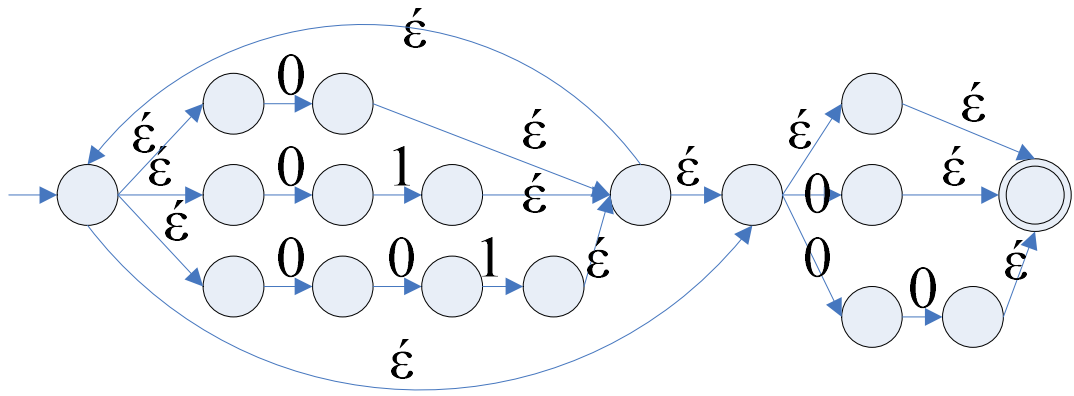
(1)



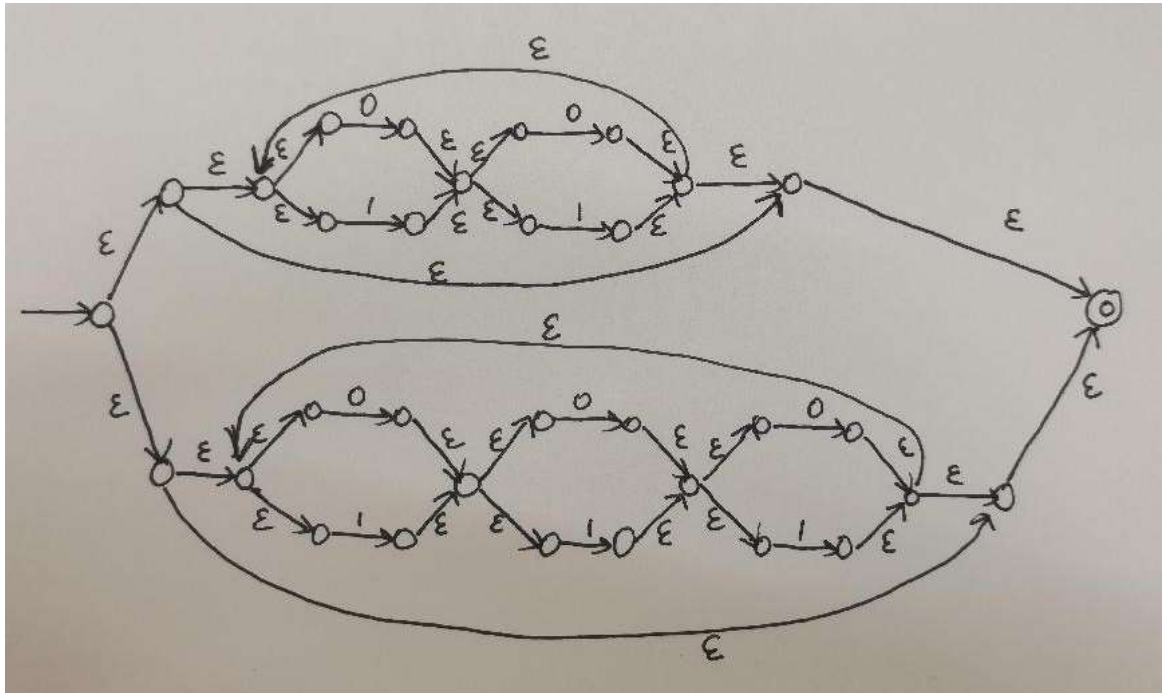
(2)



(3)



(4)



6. 构造等价于图 4-24 至图 4-27 所示 DFA 的正则表达式。

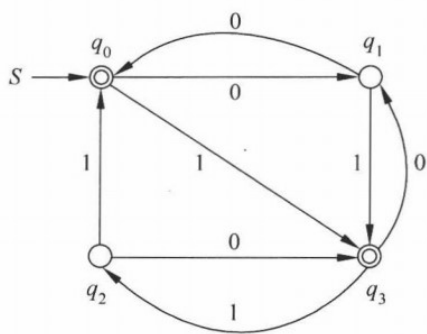


图 4-24 DFA M_1

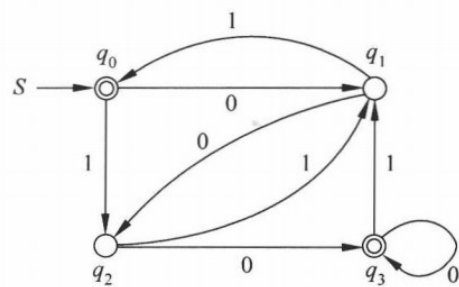
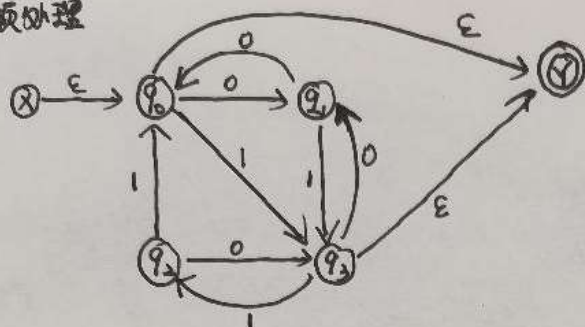


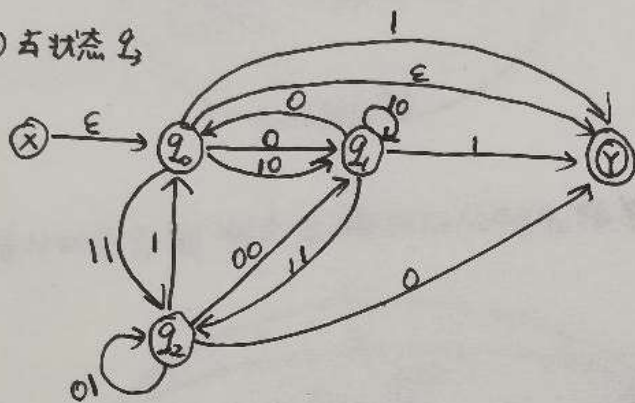
图 4-25 DFA M_2

图 4-24:

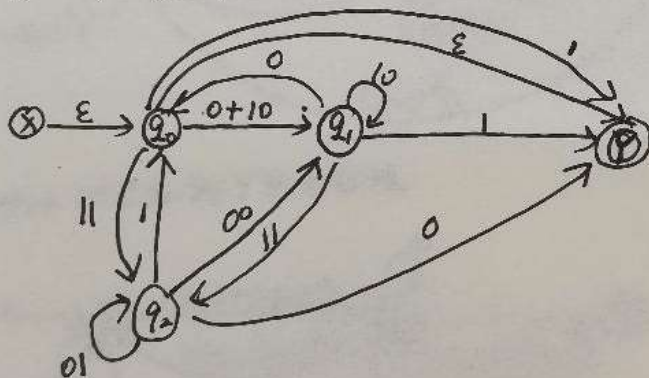
(1) 预处理



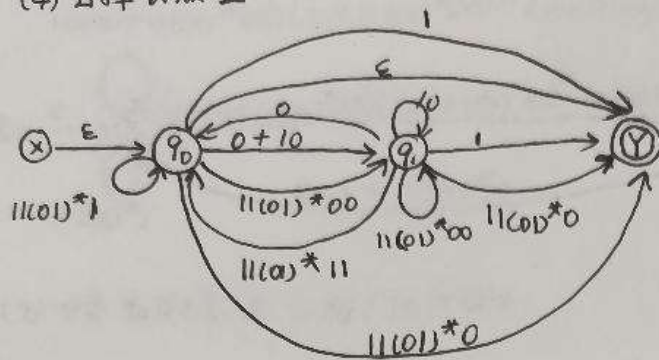
(2) 去状态 q_3



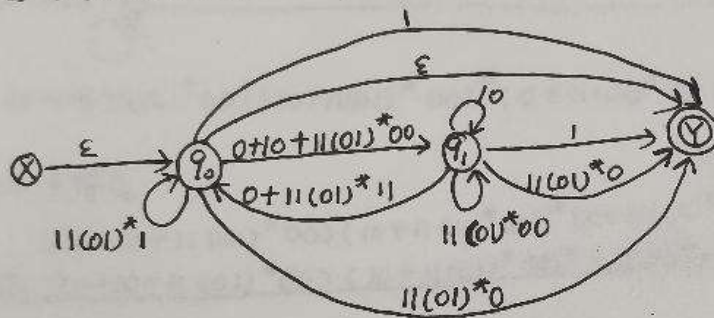
(3) 合并从 q_0 状态到 q_1 状态由两条并行弧



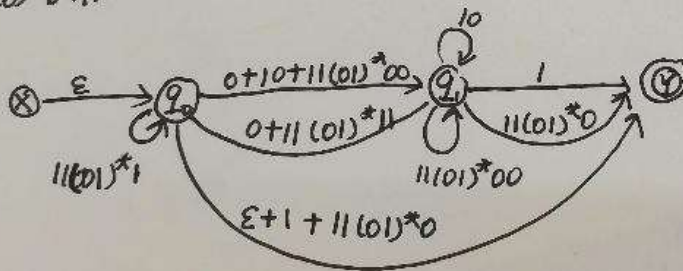
(4) 去掉状态 q_2



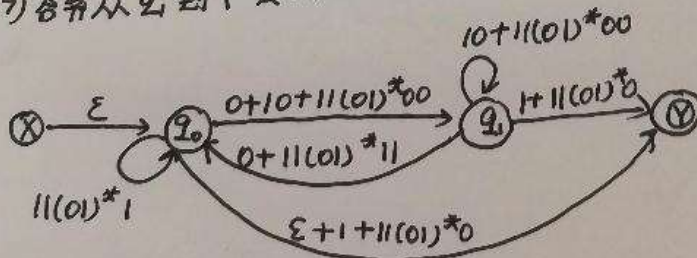
(5) 合并从状态 q_0 到状态 q_1 的及从 q_1 到 q_0 的并行弧



(6) 合并从状态 q_0 到 Y 的并行弧

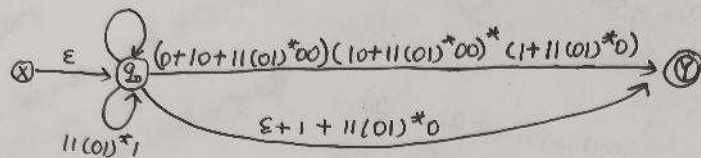


(7) 合并从 q_0 到 Y 及 q_1 到 q_0 的并行弧

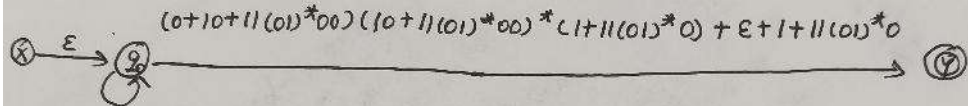


(8) 去掉状态 q_1 .

$$(10+10+11(01)^*00)(10+11(01)^*00)^*(0+11(01)^*11)$$

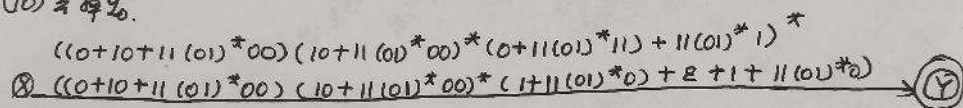


(2) 合并 q_0 到 q_0 及 q_0 到 Y 的并弧

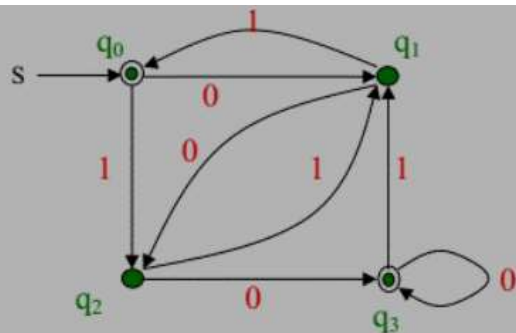


$$(0+10+11(01)^*00)(10+11(01)^*00)^*(0+11(01)^*11) + 11(01)^*1$$

(10) 去掉 q_0 .

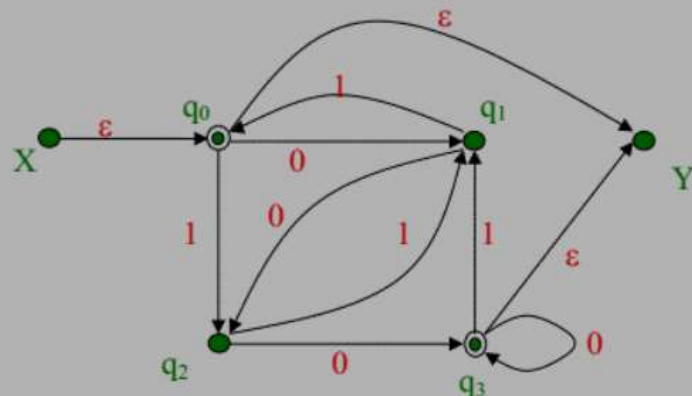


(2)

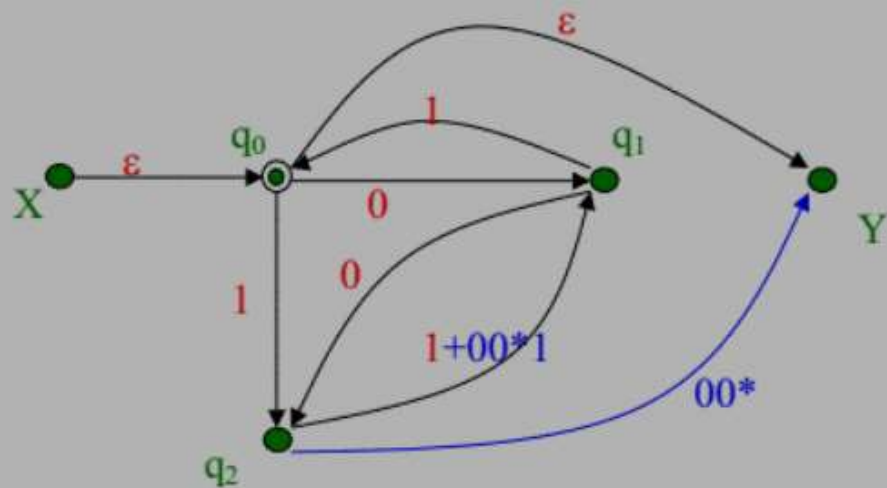


答案(之一): $(01+(1+00)((1+00*1)0)*((1+00*1)1))^*(\epsilon+(1+00)((1+00*1)0)*00^*)$

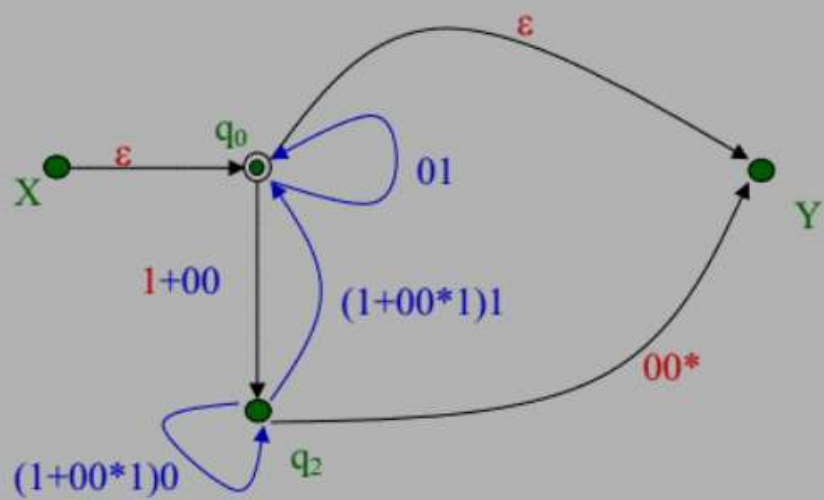
预处理:

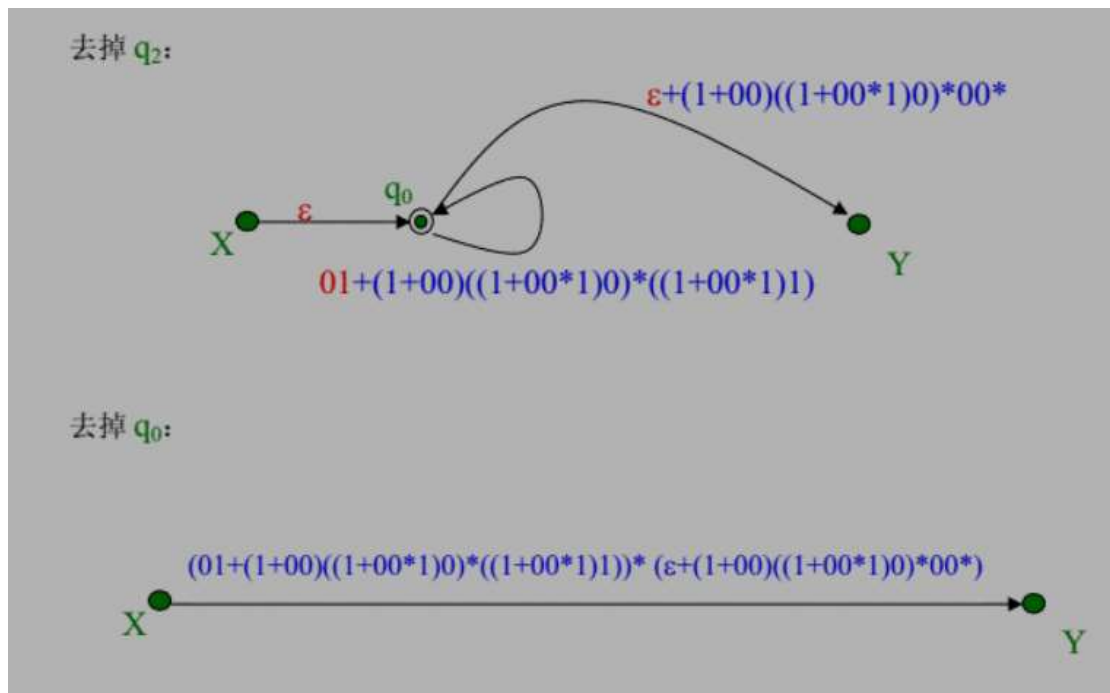


去掉 q_3 :



去掉 q_1 :





第五章习题

1. 下列语言都是字母表 $\Sigma=\{0,1\}$ 上的语言，它们哪些是正则语言，哪些不是？如果不是，请证明你的结论；如果是正则语言，请构造出它们的有穷描述(FA, RG, RE 都要写)。

(1) $\{xwx^T | x, w \in \Sigma^+\}$

是 RL。

这个语言的句子的本质特征是句子的长度大于等于 3，而且句子需要以相同的字符开头和结尾。所以，该语言的 RE 和 RG 如下：

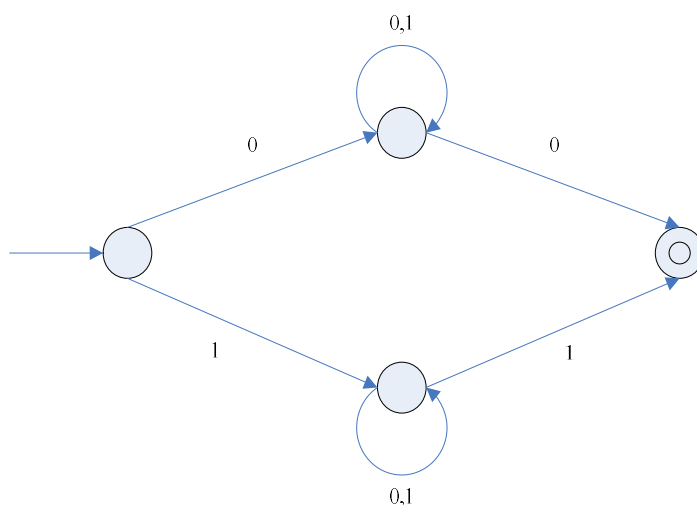
RE: $0(0+1)(0+1)^*0 + 1(0+1)^*(0+1)1$

RG: $S \rightarrow 0 A | 1 B$

$A \rightarrow 0 A | 1 A | 00 | 10$

$B \rightarrow 0 B | 1 B | 01 | 11$

该语言的 FA 如下：



(2) $\{0^{2n} | n \geq 1\}$

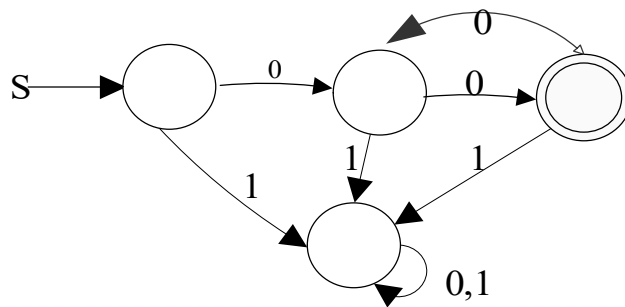
是 RL。

此语言的 RG, RE, DFA 表示。

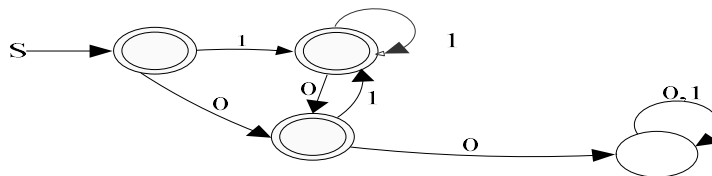
RE: $(00)(00)^*$

RG: $S \rightarrow 00 \mid 00S$

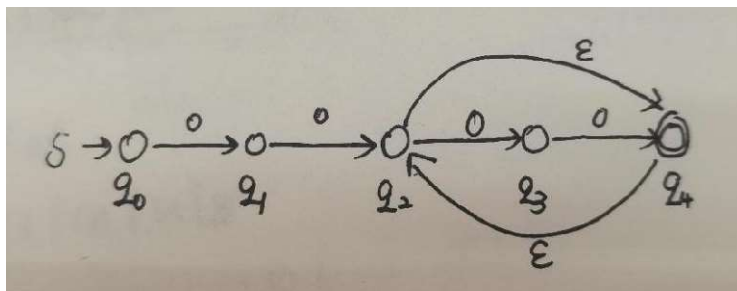
DFA:



NFA:



ϵ - NFA:



(3) $\{0^{n^2} \mid n \geq 1\}$

解答：不是RL

用泵引理证明

(1) 假设 $L = \{0^n | n \geq 1\}$ 是 RL

(2) 取 $z = 0^N, z \in L$

(3) 按照泵引理，一定存在 $z = uvw$,

取 $v = 0^k (1 \leq k \leq N)$, 此时

- $u = 0^{N-k-j}$
- $w = 0^{N^2+j-N}$

(4) 从而有,

- $uv^i w = 0^{N-k-j} (0^k)^i 0^{N^2+j-N} = 0^{N^2+(i-1)k}$

(5) 当 $i=2$ 时, 有:

- $N^2+(i-1)k = N^2+k \leq N^2+N = N(N+1) < (N+1)^2$;
- 同时 $N^2+(i-1)k = N^2+k > N^2$;
- 表明: $N^2+(i-1)k$ 不是整数的平方;

(6) 即: $i=2$ 时 $uv^i w \notin L$;

- 这与泵引理矛盾;
- 因此, L 不是 RL。

2. 用 Myhill-Nerode 定理证明第 1 题 (1) 的结论.

解: 该语言是 RL, 此时需要完整地给出 R_L 所确定的所有等价类。在给出这些等价类的过程中, 注意到该语言的句子的长度都至少为 3 是非常重要的。

R_L 将 $\{0,1\}^*$ 分成如下 7 个等价类:

$[\epsilon] = \{\epsilon\}$

$[0] = \{0\}$

$[1] = \{1\}$

$[01] = \{x | x \text{ 以 } 0 \text{ 开头, 以 } 1 \text{ 结尾, 长度至少为 } 2\} \cup \{00\}$

$[10] = \{x | x \text{ 以 } 1 \text{ 开头, 以 } 0 \text{ 结尾, 长度至少为 } 2\} \cup \{11\}$

$[00] = \{x | x \text{ 以 } 0 \text{ 开头, 以 } 0 \text{ 结尾, 长度至少为 } 3\}$

$[11] = \{x | x \text{ 以 } 1 \text{ 开头, 以 } 1 \text{ 结尾, 长度至少为 } 3\}$

R_L 的指数是有穷的, 由 Myhill-Nerode 定理, $\{xwx^T | x, w \in \Sigma^+\}$ 是 RL。

第六章习题

1. 图 1 表示的是某个 CFG 的一棵派生树, 该派生树所表达出来的派生用到了相应的文法的所有产生式。根据此派生树:

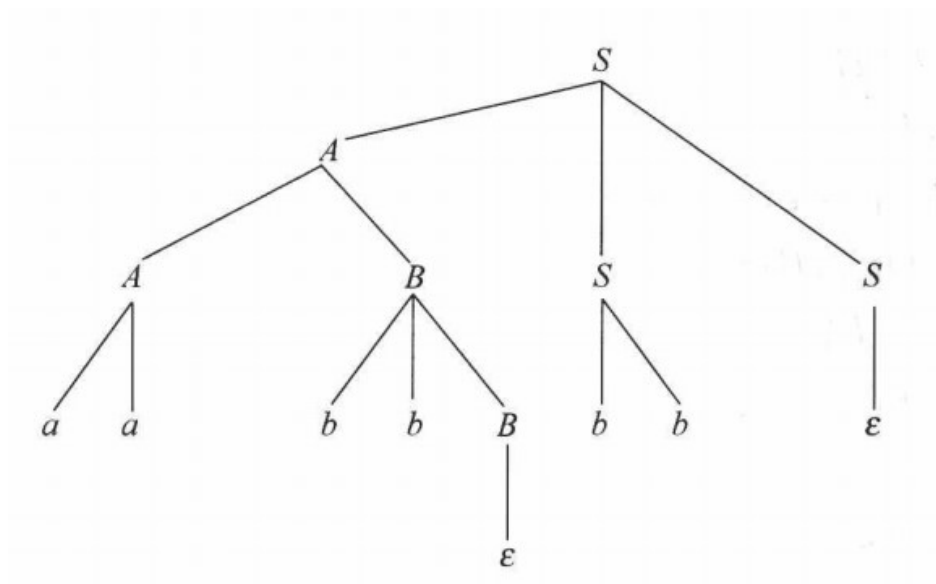


图 1

- (1) 派生树的结果的最左派生。
- (2) 写出相应的 CFG。

(1) 最左派生: $S \Rightarrow ASS$
 $\Rightarrow aaSS$
 $\Rightarrow aabBS$
 $\Rightarrow aabbbBS$
 $\Rightarrow aabbbbS$
 $\Rightarrow aabbbb$

(2) CFG
 $G = (V, T, P, S)$
 $= (\{A, B, S\}, \{a, b\},$
 $\{S \rightarrow ASS \mid bb \mid \epsilon, A \rightarrow AB \mid aa,$
 $B \rightarrow bbB \mid \epsilon\}, S)$

2. 删除下列文法中的无用符号。

- (1) $S \rightarrow AB \mid CA \mid a$
 $A \rightarrow a$
 $B \rightarrow BC \mid AB \mid DE \mid d$
 $C \rightarrow aB \mid b$
- (2) $S \rightarrow aABC \mid bCES \mid aE$
 $A \rightarrow bE \mid SCD \mid d$
 $B \rightarrow dFS \mid aBC$
 $C \rightarrow aES \mid bE$
 $D \rightarrow aAC \mid d$
 $E \rightarrow aCE \mid aCT \mid \epsilon$
 $F \rightarrow eATB \mid aF$
 $T \rightarrow eT \mid eF$

解: (1)

$S \rightarrow AB \mid CA \mid a$
 $A \rightarrow a$
 $B \rightarrow BC \mid AB \mid d$
 $C \rightarrow aB \mid b$

(2)

消除非产生的 B、F、T:

$S \rightarrow bCES \mid aE$
 $A \rightarrow bE \mid SCD \mid d$
 $C \rightarrow aES \mid bE$

$D \rightarrow aAC|d$

$E \rightarrow aCE|\epsilon$

消除非可达的 A、D

$S \rightarrow bCES|aE$

$C \rightarrow aES|bE$

$E \rightarrow aCE|\epsilon$

3. 消除下列文法中的 ϵ -产生式

(1) $S \rightarrow ABCDE|aB|\epsilon$

$A \rightarrow aBCA|BC|\epsilon$

$B \rightarrow b|bB|\epsilon$

$C \rightarrow c|cC|\epsilon$

$D \rightarrow d|dD|\epsilon$

$E \rightarrow e|eE|\epsilon$

(2) $S \rightarrow ABDC$

$A \rightarrow BD | aa|\epsilon$

$B \rightarrow aB | a$

$C \rightarrow DC|c|\epsilon$

$D \rightarrow \epsilon$

3. (1) $S \rightarrow ABCDE|aB|BCDE|ACDE|ABDE|ABCE|ABCD|$
 $CDE|BDE|BCE|BCD|ADE|ACE|ACD|ABE|ABD|ABC|$
 $DE|CE|CD|BE|BD|BC|AE|AD|Ac|AB|$
 $A|B|C|D|E|a$
 $A \rightarrow aBCA|BC|aCA|aBA|aBC|aA|aC|aB|a|C|B$
 $B \rightarrow b|bB$
 $C \rightarrow c|cC$
 $D \rightarrow d|dD$
 $E \rightarrow e|eE$

(2) $\exists D \rightarrow \epsilon: S \rightarrow ABC$

$A \rightarrow B|aa|\epsilon$

$B \rightarrow aB|a$

$C \rightarrow C|c|\epsilon$

$\exists A \rightarrow \epsilon: S \rightarrow ABC|BC$

$A \rightarrow B|aa$

$B \rightarrow aB|a$

$C \rightarrow C|c|\epsilon$

$\exists C \rightarrow \epsilon: S \rightarrow ABC|BC|AB|B$

$A \rightarrow B|aa$

$B \rightarrow aB|a$

$C \rightarrow c$

2. 设 $L = \{0^n \mid n \geq 1\}$, 试构造满足要求的文法 G 。

(1) G 是 RG。

解: $S \rightarrow 0 \mid 0S$

(2) G 是 CFG, 但不是 RG。

解: $S \rightarrow 0 \mid 0S \mid SS$

(3) G 是 CSG, 但不是 CFG。

解: $S \rightarrow 0 \mid 0S \mid AS$

$AS \rightarrow SA$

$AS \rightarrow 0A$

$0A \rightarrow S0$

$0AS \rightarrow 00$

(4) G 是短语结构文法, 但不是 CSG。

解: $S \rightarrow 0 \mid 0S \mid AS$

$AS \rightarrow SA \mid ABB$

$ABB \rightarrow AS$

$AB \rightarrow A \mid \varepsilon$

6. 设文法 G 的产生式集如下, 请给出 G 的每个语法范畴代表的集合。

$S \rightarrow aSa \mid aaSaa \mid aAa$

$A \rightarrow bA \mid bbbA \mid bB$

$B \rightarrow cB \mid cC$

$C \rightarrow ccC \mid DD$

$D \rightarrow dD \mid d$

解: 注意到 D 产生式只有 $D \rightarrow dD \mid d$, 表示从 D 出发, 可以产生任意多个 d , 且 d 的个数至少为 1。所以

$set(D) = \{d\}^+$ 。

C 产生式有两个, 一个为 $C \rightarrow ccC$, 另一个为 $C \rightarrow DD$ 。注意到产生式 $C \rightarrow ccC$ 无法使变量 C 消失, 但是, 它能够产生符号串

$c^{2n} C \quad n \geq 1$

所以, 要想使符号 C 消失, 必须使用产生式 $C \rightarrow DD$ 。再看产生式 $C \rightarrow DD$, 由于

形式语言与自动机理论教学参考书

$set(D) = \{d\}^+$, 所以根据产生式 $C \rightarrow DD$ 可以得到

$set(D)set(D) = \{d\}^+ \{d\}^+ = \{d\}\{d\}^+$

这是 C 能够推出来的一部分字符串的集合, 即它们是 $set(C)$ 的一部分元素: $\{d\}\{d\}^+$

$set(C)$ 。 $C \rightarrow ccC$ 告诉我们, $\{d\}\{d\}^+$ 的前面还可以有前缀 c^{2n} ($n \geq 1$), 故

$set(C) = \{c\}^{2n} \{d\}\{d\}^+ \quad n \geq 0$

类似地, 容易得到

$set(B) = \{c\}^+ \{d\}\{d\}^+$

$set(A) = \{b\}^+ set(B) = \{b\}^+ \{c\}^+ \{d\}\{d\}^+$

$set(S) = \{a^n b^m c^h d^k a^n \mid n \geq 1, m \geq 1, h \geq 1, k \geq 2\}$

注意:

$set(S) \neq \{a\}^+ \{b\}^+ \{c\}^+ \{d\}\{d\}^+ \{a\}^+$

8. 设 $\Sigma = \{0,1\}$, 请给出 Σ 上的下列语言的文法。

(1) 所有以 0 开头的串。

解 1: $S \rightarrow 0 \mid S0 \mid S1$ 。

S 先生成任意的 0,1 串,最后在这个 0,1 串之前生成一个 0,从而保证生成的串是以 0 开头的串。

解 2:

$S \rightarrow 0A$

$A \rightarrow 0 \mid 1 \mid 0A \mid 1A$

S 先生成 0A,然后“将任务交给变量 A ”,由 A 生成 0 后面的任意 0,1 符号串。

(2) 所有以 0 开头,以 1 结尾的串。

解 1:

$S \rightarrow 0A1$

$A \rightarrow \varepsilon \mid 0A \mid 1A$

先由 S 生成以 0 开头以 1 结尾的句型,然后由 0,1 之间的 A 生成中间部分。由于 01 本身也是满足要求的串,所以 A 可以产生 ε 。

解 2:

$S \rightarrow 0A$

$A \rightarrow 1 \mid 0A \mid 1A$

用产生式 $S \rightarrow 0A$ 保证产生的字符串是以 0 开头的,产生式 $A \rightarrow 0A \mid 1A$ 被用来生成开头字符 0 之后,结尾字符 1 之前的所有 0,1 子串,产生式 $A \rightarrow 1$ 是使除了 S 之外的所有其他句型中的惟一变量 A 变成终极符号的惟一产生式,该产生式保证了在句型的尾部生成一个 1——句子是以 1 结尾的。

(3) 所有以 11 开头,以 11 结尾的串。

解: 本题的解法与上一题类似,只不过是用 11 分别代替了字符串的首字符 0 和字符串的尾字符 1,其他位置的字符不变。因此,相应的文法如下。对比本题的解法与上题的解法,读者可以进一步理解文法是经过对语言的结构描述来定义语言的。所以,要构造一个给定语言的文法,最重要的是找出该语言的结构特征。

解 1:

$$S \rightarrow 11A11$$
$$A \rightarrow \varepsilon | 0A | 1A$$

解 2:

$$S \rightarrow 11A$$
$$A \rightarrow 11 | 0A | 1A$$

(4) 所有最多有一对连续的 0 或者最多有一对连续的 1 的串。

解: 根据题意,语言中的句子 x 可以有如下几种情况:

- ① x 中既没有成对的 0,也没有成对的 1;
- ② x 恰有一对连续的 0;
- ③ x 恰有一对连续的 1;
- ④ x 恰有一对连续的 0 且恰有一对连续的 1。

对以上 4 类句子,可以依次分别用 A, B, C, D 表示:

$$S \rightarrow A | B | C | D$$
$$A \rightarrow \varepsilon | A' | A''$$
$$A' \rightarrow 0 | 01 | 01A'$$
$$A'' \rightarrow 1 | 10 | 10A''$$
$$B \rightarrow B'00B''$$
$$B' \rightarrow 1 | 01 | 1B' | 01B'$$
$$B'' \rightarrow 1 | 10 | 1B'' | 10B''$$
$$C \rightarrow C'11C''$$
$$C' \rightarrow 0 | 10 | 0C' | 10C'$$
$$C'' \rightarrow 0 | 01 | 0C'' | 01C''$$
$$D \rightarrow E00F11H | 11G00K$$
$$E \rightarrow 1 | 1E' | E'$$
$$E' \rightarrow 01E' | 01$$
$$H \rightarrow 0 | H0 | H'$$

$$\begin{aligned}
I &\rightarrow 0|0I'|I' \\
I' &\rightarrow 10I'|10 \\
K &\rightarrow 1|K'1|K' \\
K' &\rightarrow 10|10K'
\end{aligned}$$

(5) 所有最多有一对连续的 0 并且最多有一对连续的 1 的串。

解 1: 首先要弄清楚 L 包含的串, 然后再设计文法。实际上, L 中的串有如下几种形式:

- ① 既不含成对的 0, 也不含成对的 1。
- ② 只含惟一的一对 0。
- ③ 只含惟一的一对 1。
- ④ 含有一对 0 和一对 1, 这一对 0 在这一对 1 的前面。
- ⑤ 含有一对 0 和一对 1, 这一对 0 在这一对 1 的后面。

所以, 我们可以将 L 写出如下形式:

$$\begin{aligned}
L = \{ &x \mid x \text{ 中不含连续的 0, 也不含连续的 1} \} \\
&\cup \{ x00y \mid x, y \text{ 是不含连续的 0 也不含连续的 1 的串, 并且 } x \text{ 不以 0 结尾, } y \text{ 不以 0 开头} \} \\
&\cup \{ x11y \mid x, y \text{ 是不含连续的 0 也不含连续的 1 的串, 并且 } x \text{ 不以 1 结尾, } y \text{ 不以 1 开头} \} \\
&\cup \{ x00y11z \mid x, y, z \text{ 是不含连续的 0 也不含连续的 1 的串, 并且 } x \text{ 不以 0 结尾, } y \text{ 不以 0 开头, } y \text{ 不以 1 结尾, } z \text{ 不以 1 开头} \} \\
&\cup \{ x11y00z \mid x, y, z \text{ 是不含连续的 0 也不含连续的 1 的串, 并且 } x \text{ 不以 1 结尾, } y \text{ 不以 1 开头, } y \text{ 不以 0 结尾, } z \text{ 不以 0 开头} \} .
\end{aligned}$$

根据上述几种模式, 进行文法设计。

$$\begin{aligned}
S &\rightarrow 00|11|0011|1100|E|A00B|D11C|A00F11C|D11G00B \\
E &\rightarrow \varepsilon|E'E'' \\
E' &\rightarrow 0|01|01E' \\
E'' &\rightarrow 1|10|10E'' \\
A &\rightarrow 1|1A'|A' \\
A' &\rightarrow 01A'|01 \\
B &\rightarrow 1|B'1|B' \\
B' &\rightarrow 10|10B'
\end{aligned}$$

$$\begin{aligned}
 E &\rightarrow \varepsilon \mid E' \mid E'' \\
 E' &\rightarrow 0 \mid 01 \mid 01 E' \\
 E'' &\rightarrow 1 \mid 10 \mid 10 E'' \\
 A &\rightarrow 1 \mid 1 A' \mid A' \\
 A' &\rightarrow 01 A' \mid 01 \\
 B &\rightarrow 1 \mid B' 1 \mid B' \\
 B' &\rightarrow 10 \mid 10 B' \\
 D &\rightarrow 0 \mid 0 D' \mid D' \\
 D' &\rightarrow 10 D' \mid 10
 \end{aligned}$$

文 法 第
2 章

$$\begin{aligned}
 C &\rightarrow 0 \mid C 0 \mid C' \\
 C' &\rightarrow 01 \mid 01 C' \\
 F &\rightarrow 10 \mid 10 F \\
 G &\rightarrow 01 \mid 01 G
 \end{aligned}$$

解 2: 也可以根据上述分析, 得出以下正则文法(RG)。

$$\begin{aligned}
 S &\rightarrow \varepsilon \mid 0 \mid 1 \mid 0 A \mid 1 B \\
 A &\rightarrow 0 \mid 1 \mid 0 C \mid 1 B \\
 B &\rightarrow 0 \mid 1 \mid 0 A \mid 1 G \\
 C &\rightarrow 1 \mid 1 D \\
 D &\rightarrow 0 \mid 1 \mid 0 C \mid 1 E \\
 E &\rightarrow 0 \mid 0 F \\
 F &\rightarrow 1 \mid 1 E \\
 G &\rightarrow 0 \mid 0 H \\
 H &\rightarrow 0 \mid 1 \mid 0 I \mid 1 G \\
 I &\rightarrow 1 \mid 1 J \\
 J &\rightarrow 0 \mid 0 I
 \end{aligned}$$

(6) 所有长度为偶数的串。

解 1: 用两个变量 A 和 B 交替, 以保证长度为偶数。

$$S \rightarrow \varepsilon | 0A | 1A$$
$$A \rightarrow 0 | 1 | 0B | 1B$$
$$B \rightarrow 0A | 1A$$

解 2: 按照偶数长的字符串可以被切分成若干个长度为 2 的子串的特性进行构造, 让文法每次长出两个终极符号, 这两个终极符号可以是 00, 01, 10, 或者 11。

$$S \rightarrow \varepsilon | 00S | 01S | 10S | 11S$$

(7) 所有包含子串 01011 的串。

解:

$$S \rightarrow A01011A$$
$$A \rightarrow \varepsilon | 0A | 1A$$

(8) 所有含有 3 个连续 0 的串。

解:

$$S \rightarrow 0A | 1S$$
$$A \rightarrow 0B | 1S$$
$$B \rightarrow 0 | 0C | 1S$$

形式语言与自动机理论教学参考书

$$C \rightarrow 0 | 1 | 0C | 1C$$

9. 设 $\Sigma = \{a, b, c\}$, 构造下列语言的文法。

(1) $L_1 = \{a^n b^n \mid n \geq 0\}$ 。

解: $S \rightarrow ab \mid aSb$

(2) $L_2 = \{a^n b^m \mid n, m \geq 1\}$ 。

解 1:

$S \rightarrow aS \mid aA$

$A \rightarrow b \mid bA$

解 2:

$S \rightarrow AB$

$A \rightarrow a \mid aA$

$B \rightarrow b \mid bB$

(3) $L_3 = \{a^n b^n a^n \mid n \geq 1\}$ 。

解: 参考例 2-15 题, 可以得到如下文法:

$S \rightarrow aBC \mid aSBC,$

$CB \rightarrow BC$

$aB \rightarrow ab$

$bB \rightarrow bb$

$bC \rightarrow ba$

$aC \rightarrow aa$

根据此文法, 可以做进一步的优化, 优化后可得:

$S \rightarrow aba \mid aSBa$

$bB \rightarrow bb$

$aB \rightarrow Ba$

(4) $L_4 = \{a^n b^m a^k \mid n, m, k \geq 1\}$ 。

解 1:

$S \rightarrow aS \mid aA$

$A \rightarrow bB \mid bA$

$B \rightarrow a \mid aB$

(5) $L_5 = \{ awa \mid a \in \Sigma, w \in \Sigma^+ \}$ 。

解 1:

$S \rightarrow aAa$

$A \rightarrow a \mid b \mid c \mid aA \mid bA \mid cA$

解 2: 构造成正则文法 (RG):

$S \rightarrow aA$

$A \rightarrow aB \mid bB \mid cB$

$B \rightarrow a \mid aB \mid bB \mid cB$

(6) $L_6 = \{ xwx^T \mid x, w \in \Sigma^+ \}$ 。

解 1:

$S \rightarrow aAa \mid bAb \mid cAc$

$A \rightarrow a \mid b \mid c \mid aA \mid bA \mid cA$

解 2:

$S \rightarrow aA \mid bB \mid cC$

$A \rightarrow aA' \mid bA' \mid cA'$

$A' \rightarrow bA'' \mid cA'' \mid aA''' \mid a$

$A'' \rightarrow a \mid aA''' \mid bA''' \mid cA'''$

$B \rightarrow aB' \mid bB' \mid cB'$

$B' \rightarrow aB'' \mid cB'' \mid bB''' \mid b$

$B'' \rightarrow b \mid bB''' \mid aB''' \mid cB'''$

$C \rightarrow aC' \mid bC' \mid cC'$

$C' \rightarrow aC'' \mid bC'' \mid cC''' \mid c$

$C'' \rightarrow c \mid cC''' \mid aC''' \mid bC'''$

(8) $L_8 = \{ xx^T w \mid x, w \in \Sigma^+ \}$ 。

解:

$S \rightarrow DT$

$D \rightarrow aa \mid bb \mid cc \mid aDa \mid bDb \mid cDc$

$T \rightarrow a \mid b \mid c \mid aD \mid bD \mid cD$

2. 构造识别下列语言的 DFA(给出相应 DFA 的形式描述或者画出它们的状态转移图)。

(2) $\{0,1\}^*$

解:本题的关键是保证接受的串的长度至少为 1,相应的 DFA 如图 3-18(a)所示。



图 3-18(a) 题 2(2)的 DFA

(3) $\{x \mid x \in \{0,1\}^* \text{ 且 } x \text{ 中不含形如 } 00 \text{ 的子串}\}$

解:构造要点:自动机启动并读入一个字符后,就将“精力”集中在考察是否出现 00 子串上,一旦发现子串 00,就进入陷阱状态。本小题的构造结果如图 3-18(b)所示。

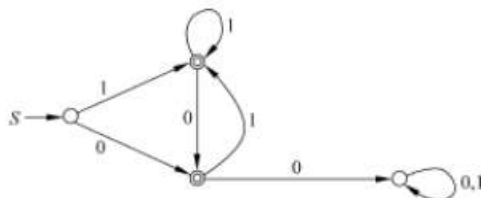


图 3-18(b) 题 2(3)的 DFA

(7) $\{x \mid x \in \{0,1\}^* \text{ 且当把 } x \text{ 看成二进制数时, } x \text{ 模 } 5 \text{ 与 } 3 \text{ 同余,要求当 } x \text{ 为 } 0 \text{ 时, } |x| = 1, \text{ 且当 } x \neq 0 \text{ 时, } x \text{ 的首字符为 } 1\}.$

解:构造要点:

① 以 0 开头的串都不能被此 DFA 接受,包括字符串 0。所以,如果 DFA 在启动状态读入的符号为 0,则直接进入陷阱状态。

② 该 DFA 共有 7 个状态:开始状态、陷阱状态、终止状态各一个,在相应的状态转移图中,终止状态和其余 4 种状态构成最大的强连通子图。

③ 其余部分请参考主教材例 3-5。

(9) $\{x \mid x \in \{0,1\}^* \text{ 且 } x \text{ 以 } 0 \text{ 开头以 } 1 \text{ 结尾}\}.$

解:构造要点:

① 启动时,只要考虑开始符号是否为 0。

形式语言与自动机理论教学参考书

② 以 1 开头的字符串都是不可接受的。

③ 在读入一个字符 0 之后,当读到 1 时,可将这个 1 先当成结尾的 1,如果是,就停止并接受;如果不是,就继续向前扫描。

④ 被接受串的长度至少为 2。

本小题的构造结果如图 3-18(c)所示。

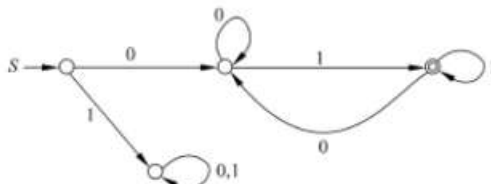


图 3-18(c) 题 2(9)的 DFA

(11) $\{x \mid x \in \{0,1\}^* \text{ 且如果 } x \text{ 以 } 1 \text{ 结尾, 则它的长度为偶数; 如果 } x \text{ 以 } 0 \text{ 结尾, 则它的长度为奇数}\}$ 。

解: 构造要点:

① 语言根据句子长度的奇偶性对句子的结尾符号提出要求, 因此, 它将 $\{0,1\}^*$ 中的字符串分成 4 个等价类, 这 4 个等价类依次为:

[奇 0]—— $\{x \mid x \in \{0,1\}^* \text{ 不仅长度为奇数, 而且以 } 0 \text{ 结尾}\}$ 。

[奇 1]—— $\{x \mid x \in \{0,1\}^* \text{ 不仅长度为奇数, 而且以 } 1 \text{ 结尾}\}$ 。

[偶 0]—— $\{x \mid x \in \{0,1\}^* \text{ 不仅长度为偶数, 而且以 } 0 \text{ 结尾}\}$ 。

[偶 1]—— $\{x \mid x \in \{0,1\}^* \text{ 不仅长度为偶数, 而且以 } 1 \text{ 结尾}\}$ 。

可以用[奇 0]、[奇 1]、[偶 0]、[偶 1]分别标示这 4 个等价类对应的状态。显然, [奇 0]、[偶 1]对应的状态为终止状态。

② ε 不属于上述任何一个等价类, 所以, 它自己独立地构成一个等价类, 而且它是语言的句子, 该等价类对应的状态为终止状态。可以用 $[\varepsilon]$ 表示此等价类及其对应的状态。

③ $M = (\{[\varepsilon], [\text{奇 } 0], [\text{奇 } 1], [\text{偶 } 0], [\text{偶 } 1]\}, \{0,1\}, \delta, [\varepsilon], \{[\varepsilon], [\text{奇 } 0], [\text{偶 } 1]\})$, 其中 δ 由下面状态转移表表示。

	$[\varepsilon]$	[奇 0]	[奇 1]	[偶 0]	[偶 1]
0	[奇 0]	[偶 0]	[偶 0]	[奇 0]	[奇 0]
1	[奇 1]	[偶 1]	[偶 1]	[奇 1]	[奇 1]

如果我们将长度 0 也看成偶数长, 从此表可以明显地看出, 在读入一个字符时, 已处理的输入串的相应前缀长度的奇偶性“交替变化”: 奇变偶, 偶变奇; 而 0,1 区分正好体现了读进来的符号: 是 0 就是 0, 是 1 就是 1。

④ 如果认为 ε 不是语言的句子, 则只要将 $[\varepsilon]$ 从终止状态集中删除即可。

8. 证明: 对于任意的 DFA $M_1 = (Q, \Sigma, \delta, q, F_1)$, 存在 DFA $M_2 = (Q, \Sigma, \delta, q, F_2)$, 使得 $L(M_2) = \Sigma^* - L(M_1)$ 。

证明:

① 构造 M_2 。

设 DFA $M_1 = (Q, \Sigma, \delta, q, F_1)$ 。

取 DFA $M_2 = (Q, \Sigma, \delta, q, Q - F_1)$ 。

② 证明 $L(M_2) = \Sigma^* - L(M_1)$

对任意 $x \in \Sigma^*$,

$x \in L(M_2) = \Sigma^* - L(M_1) \iff \delta(q, x) \in Q - F_1 \iff \delta(q, x) \in Q \text{ 并且 } \delta(q, x) \notin F_1 \iff x \in \Sigma^* \text{ 并且 } x \notin L(M_1) \iff x \in \Sigma^* - L(M_1)$

③ 对于任意的 DFA $M_1 = (Q, \Sigma, \delta, q, F_1)$, 构造 DFA $M_2 = (Q, \Sigma, \delta, q, Q - F_1)$

9. 对于任意的 DFA $M_1 = (Q, \Sigma, \delta, q_0, F_1)$, 请构造 DFA $M_2 = (Q, \Sigma, \delta, q_0, F_2)$, 使得 $L(M_2) = L(M_1)^r$ 。其中, $L(M)^r = \{x \mid x^r \in L(M)\}$ 。

证明:

① 构造 ε -NFA M 使得 $L(M) = L(M_1)^r$ 。

设 DFA $M_1 = (Q, \Sigma, \delta, q_0, F_1)$,

先取 ε -NFA $M = (Q, \Sigma, \delta, q, \{q_0\})$, 其中

$Q = Q \cup \{q\}, q \notin Q$

对于任意的 $q, p \in Q, a \in \Sigma$,

$q \in \delta(p, a) \quad \delta(q, a) = p$

$\delta(q, \varepsilon) = F_1$

② 证明 $L(M) = L(M_1)^r$ 。

对任意 $a_1 a_2 \cdots a_n \in \Sigma^+$,

$a_1 a_2 \cdots a_n \in L(M) \quad q \xrightarrow{a_1} a_2 \cdots a_n \xrightarrow{a_n} q_f \xrightarrow{\varepsilon} F_1$
 $\vdash a_1 a_2 \cdots q_{n-1} a_n \xrightarrow{a_n} a_1 a_2 \cdots q_0 \xrightarrow{\varepsilon} F_1$ 并且 $q_f \in F_1$

$q_f \in \delta(q, \varepsilon), q \in \delta(q_f, a_1), q_1 \in \delta(q, a_2), \cdots, q_{n-1} \in \delta(q_{n-2}, a_n)$, 并且 $q_f \in F_1$

$\delta(q_f, a_1) = q_f, \delta(q, a_2) = q_1, \cdots, \delta(q_{n-1}, a_n) = q_{n-1}$, 并且 $q_f \in F_1$

$q_1 a_n a_{n-1} \cdots a_1 \vdash a_1 q_{n-1} a_{n-1} \cdots a_1 \vdash \cdots \vdash a_n a_{n-1} \cdots q_2 a_2 a_1 \vdash a_n a_{n-1} \cdots a_1 q_f$
 $\vdash a_n a_{n-1} \cdots a_1 a_1 q_f$

$a_n a_{n-1} \cdots a_1 a_1 x \in L(M_1)$

③ 按照将 ε -NFA 转换成等价的 NFA, 再将 NFA 转换成等价的 DFA 的方法, 将此 ε -NFA 转换成满足要求的 DFA $M_2 = (Q, \Sigma, \delta, q_0, F_2)$ 。

10. 构造识别下列语言的 NFA。

(2) $\{x \mid x \in \{0, 1\}^+ \text{ 并且 } x \text{ 中含形如 } 10110 \text{ 的子串}\}$ 。

解: 本小题的构造结果如图 3-19(a)所示。



图 3-19(a) 题 10(2)的 NFA

有穷状态自动机

第
3
章

(3) $\{x \mid x \in \{0, 1\}^+ \text{ 并且 } x \text{ 中不含形如 } 10110 \text{ 的子串}\}$ 。

解: 构造要点:

① 虽然在不考虑空串的前提下, 本小题所给的语言为上面的第(2)小题所给语言的补, 决不能通过将上题所给出的 NFA 的终止状态和非终止状态互换来构造本题要求的 NFA。也就是说, 习题 8 所给的方法对 NFA 来说, 是不适应的。

② 需要构造接收相应语言的 DFA——一种特殊的 NFA。具体如图 3-19(b)所示。

③ 先根据习题 8 和上一小题的结果构造主题框架, 然后逐渐补充。

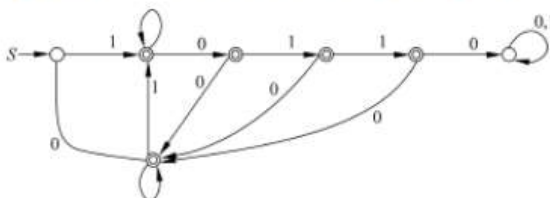


图 3-19(b) 题 10(3)的 NFA

(5) $\{x \mid x \in \{0,1\}^* \text{ 并且 } x \text{ 以 } 0 \text{ 开头以 } 1 \text{ 结尾}\}$ 。

解：具体构造结果如图 3-19(c)所示。

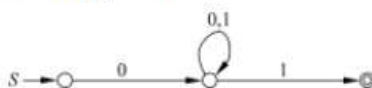


图 3-19(c) 题 10(5)的 NFA

(7) $\{x \mid x \in \{0,1\}^* \text{ 且如果 } x \text{ 以 } 1 \text{ 结尾, 则它的长度为偶数; 如果 } x \text{ 以 } 0 \text{ 结尾, 则它的长度为奇数}\}$ 。

解：具体构造结果如图 3-19(d)所示。

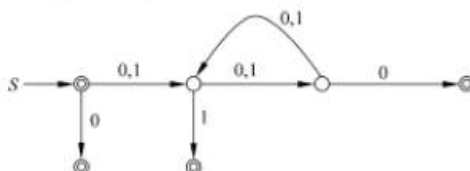


图 3-19(d) 题 10(7)的 NFA

如果认为 ε 不是语言的句子, 则相应的 NFA 的状态转移图如图 3-19(e)所示。

形式语言与自动机理论教学参考书

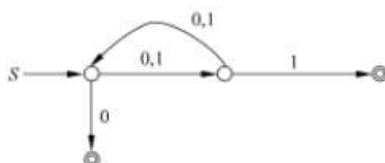


图 3-19(e) 题 10(7) ε 不是语言句子时的 NFA

(8) $\{x \mid x \in \{0,1\}^+ \text{ 且 } x \text{ 的首字符和尾字符相等}\}$ 。

解：具体构造结果如图 3-19(f)所示。

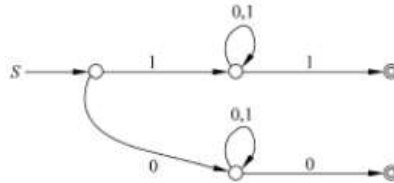


图 3-19(f) 题 10(8)的 NFA

12. 证明:对于任意的 NFA,存在与之等价的 NFA,该 NFA 最多只有一个终止状态。

证明提示:一般地,NFA 会有多个状态,当该 NFA 的可达的终止状态数小于等于 1 时,删除所有的不可达状态即可。否则,先删除所有的不可达状态,再构造与之等价的只有一个终止状态的 NFA。最简单的办法是对不含不可达状态的 NFA,将原来的终止状态均改为非终止状态,然后新增加一个终止状态,并将原来存在的到达原来终止状态弧进行复制,使之也可以到达新的终止状态。

14. 根据上面相关习题的结果,构造识别下列语言的 ϵ -NFA。

(5) $\{x \mid x \in \{0,1\}^+ \text{ 且 } x \text{ 中不含形如 } 00 \text{ 的子串}\} \cap \{x \mid x \in \{0,1\}^+ \text{ 且 } x \text{ 中不含形如 } 11 \text{ 的子串}\}$ 。

解:根据习题 2 的第(4)小题的结果进行改造,将接受以 1 结尾的串的终止状态到自身的标记为 1 的弧引向陷阱状态即可,结果如图 3-19(g)所示。

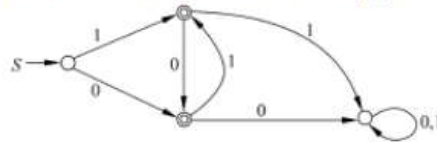


图 3-19(g) 题 14(5)的 ϵ -NFA

20. 图 3-20 所给 DFA 对应的右线性文法。

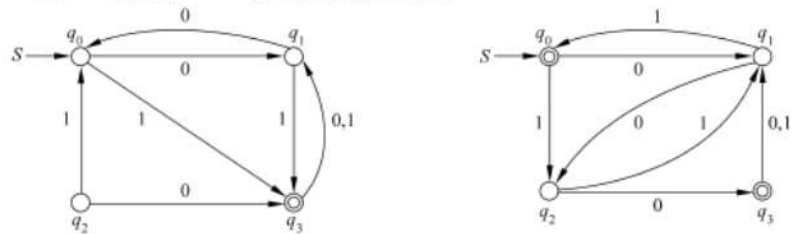


图 3-20 两个不同的 DFA

形式语言与自动机理论教学参考书

解:

G_1 对应图 3-20 中左边的 DFA:

删除不可达状态 q_2 , 让 S 对应于 q_0 , A 对应于 q_1 , B 对应于 q_3 。

$S \rightarrow 1|0A|1B$

$A \rightarrow 1|0S|1B$

$B \rightarrow 0A|1A$

G_2 对应图 3-20 中右边的 DFA:

让 S 对应于 q_0 , A 对应于 q_1 , B 对应于 q_2 , C 对应于 q_3 。

$S \rightarrow \epsilon|0A|1B$

$A \rightarrow 1|1S|0B$

$B \rightarrow 0|0C|1A$

$C \rightarrow 0A|1A$

2. 理解如下正则表达式, 说明它们表示的语言。

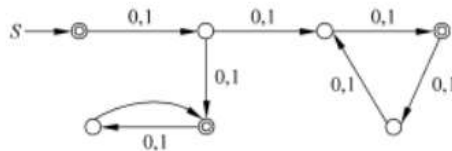
说明: 这一类题的目的在于使读者能读懂正则表达式, 对不同的人, 用自然语言描述是不尽相同的, 基本要求是说出相应语言的主要特点。例如:

$(00 + 11)^*$ 表示的语言的主要特征是 0 和 1 都是各自成对出现的。

$(1 + 01 + 001)^* (\epsilon + 0 + 00)$ 表示的语言的主要特征是不含连续的 3 个 0。

$((0 + 1)(0 + 1))^* ((0 + 1)(0 + 1)(0 + 1))^*$ 表示所有长度为 $3 \times n + 2 \times m$ 的 0, 1 串 ($n \geq 0, m \geq 0$)。

例如: 对 $((0 + 1)(0 + 1))^* + ((0 + 1)(0 + 1)(0 + 1))^*$, 它表示的是长度为偶数的 0, 1 串和长度是 3 的倍数的 0, 1 串构成的集合。所以, 相应的 FA 如图 4-28 所示。



1. 下列语言都是字母表 $\Sigma = \{0,1\}$ 上的语言, 它们哪些是 RL, 哪些不是 RL? 如果不是 RL, 请用证明你的结论; 如果是 RL, 请构造出它们的有穷描述 (FA, RG, 或者 RE)。

(1) $\{0^{2^n} \mid n \geq 1\}$ 。

解: 是 RL。

有些初学者可能会错误地认为该语言不是 RL, 并且用关于 RL 的泵引理证明该语言不是 RL。此时就应该仔细地检查一下, 看在证明的什么地方出了错误, 这个错误很可能就是你对泵引理不理解的地方。实际上, 很容易给出此语言的 RG, RE, DFA 表示。

RE: $(00)^+(00)^*$

RG: $S \rightarrow 00 \mid 00S$

(2) $\{0^{n^2} \mid n \geq 1\}$ 。

解: 不是 RL。

假设 $\{0^{n^2} \mid n \geq 1\}$ 是 RL, N 是 RL 的泵引理所说的正整数, 取 $z = 0^{N^2}$, $v = 0^i$, $0 < k \leq N$, 此时

$$uv^i w = 0^{N^2 - k + ik} = 0^{N^2 + (i-1)k}$$

取 $i = 2$, 并注意 $0 < k \leq N$, 可以得到

$$N^2 < N^2 + (2-1)k = N^2 + k \leq N^2 + N = N(N+1) < (N+1)^2$$

这表明, $N^2 + (2-1)k = N^2 + k$ 不是某个整数的平方, 所以

$$uv^2 w \notin \{0^{n^2} \mid n \geq 1\}$$

与泵引理矛盾, 故 $\{0^{n^2} \mid n \geq 1\}$ 不是 RL。

(8) $\{xwx^T \mid x, w \in \Sigma^*\}$ 。

(8) $\{xwx^T \mid x, w \in \Sigma^*\}$ 。

解: 是 RL。

分析: 这个语言的句子的本质特征是句子的长度大于等于 3, 而且句子需要以相同的字符开头和结尾。所以, 很容易得到表示该语言的正则表达式和正则文法。

RE: $0(0+1)(0+1)^*0+1(0+1)^*(0+1)1$

RG: $S \rightarrow 0A \mid 1B$

$A \rightarrow 0A \mid 1A \mid 00 \mid 10$

$B \rightarrow 0B \mid 1B \mid 01 \mid 11$

$$(9) \{xx^T w \mid x, w \in \Sigma^*\}.$$

解：此语言不是 RL。

这个语言似乎与(8)题所给的语言有一点类似,但它们却属于不同的语言类。实际上该语言为非正则语言的上下文无关语言。可以用关于 RL 的泵引理证明它不是 RL。这里的关键是找到一个供讨论的句子 z 。

注意到句子的形式是 $xx^T w$, 所以, 要想在“泵”的过程中找出矛盾, 就应该不允许 z 的头两个字符是相同的, 而且还要考虑到 z 的长度大于等于 N , 和 x 与 x^T 的对应, 所以, 可以用

$$z = (01)^N (10)^N 1$$

进行相应的讨论。讨论中,取 $v = (01)^i$, 当 $i > 1$ 时,很容易找出矛盾。但是,如果取 v 为 ε 的首字符 0,将不会出现与引理矛盾的现象。

一般地,取

$$z = a_1 a_2 \cdots a_N a_N a_{N-1} \cdots a_2 a_1 a_X$$

此时,只要令 $v = a_n$,我们就无法找出矛盾来。这表明,关于 RL 的基本泵引理对此语言的证明是无效的。

难道此语言是 RL? 在前面, 前面反复强调过, 泵引理不能用来证明 L 是 RL。实际上, 从 z 的形式可以看出, 要想识别这类句子, “DFA” 需要记着目前已经读入了多少个 01, 以便后面进行匹配, 显然有无穷多种情况需要记忆, 这与 DFA 的有穷记忆功能是矛盾的。此分析表明, 该语言不可能是 RL。实际上, 可以使用第 2 题将证明的关于 RL 的扩充泵引理完成相应的证明。

此外,根据对 Σ 的分析,容易找到 Σ^* 关于 R_L 的无穷多个等价类, R_L 的指数无穷这一事实表明 L 不是 RL。

除了取 $z = (01)^N (10)^N 1$ 之外, 还可以取其他形式的 z 来完成相应的证明。

6. 设字母表 $\{0,1\}$ 上的语言 $L = \{x \mid x \text{ 中 } 1 \text{ 的个数恰好是 } 0 \text{ 的个数的二倍}\}$, 求 R_L 的所有等价类。

解:注意到 L 的句子中 0 和 1 的前后顺序是没有限制的,所以,在进行等价类的划分时,只用考虑符号串中 0 的个数和 1 的个数。这里要注意的问题是不要将 0 的个数比 1 的个数多的串全部都划入同一个等价类中。相应的等价类有,

$[0] = \{x \mid x \text{ 为 } 0, 1 \text{ 组成的串, 且 } x \text{ 中 } 0 \text{ 的个数减去 } 1 \text{ 的个数之差为 } 0\}$

空串包括在这个等价类中。

$[1] = \{x \mid x \text{ 为 } 0,1 \text{ 组成的串, 且 } x \text{ 中 } 0 \text{ 的个数减去 } 1 \text{ 的个数之差为 } 1\}$

$$[-1] = \{x \mid x \text{ 为 } 0,1 \text{ 组成的串, 且 } x \text{ 中 } 0 \text{ 的个数减去 } 1 \text{ 的个数之差为 } -1\}$$

[2] = { x | x 为 0,1 组成的串, 且 x 中 0 的个数减去 1 的个数之差为 2}

$$[-2] = \{x \mid x \text{ 为 } 0, 1 \text{ 组成的串, 且 } x \text{ 中 } 0 \text{ 的个数减去 } 1 \text{ 的个数之差为 } -2\}$$

...

$[n] = \{x \mid x \text{ 为 } 0, 1 \text{ 组成的串, 且 } x \text{ 中 } 0 \text{ 的个数减去 } 1 \text{ 的个数之差为 } n\}$

$$[-n] = \{x \mid x \text{ 为 } 0,1 \text{ 组成的串, 且 } x \text{ 中 } 0 \text{ 的个数减去 } 1 \text{ 的个数之差为 } -n\}$$

7. 用 Myhill-Nerode 定理证明第 1 题中各题的结论。

用 Myhill-Nerode 定理证明一个语言是否为 RL, 关键是证明 R_L 的指数是有穷还是无穷, 也就是说, 它将语言所在的字母表的克林闭包是分成有穷多个等价类还是分成无穷多个等价类。当该语言是 RL 时, 我们需要清楚地给出每一个等价类; 当一个语言不是 RL 时, 我们无需找出所有的等价类, 也无需说明每个等价类的全部内容, 只要说明有无穷多个等价类就可以了。

(3) $\{0^n 1^m 2^n \mid n, m \geq 1\}$ 。

解: 该语言不是 RL。因为通过 0 和 00 不在同一个等价类我们可以推测出 0^n 与 0^m 在 n 和 m 不相等的时候不会在同一个等价类中。事实上, 当 $n \neq m$ 时,

$$0^n 12^n \in \{0^n 1^m 2^n \mid n, m \geq 1\}$$

但是

$$0^m 12^n \notin \{0^n 1^m 2^n \mid n, m \geq 1\}$$

这表明, R_L 的指数是无穷的。所以, $\{0^n 1^m 2^n \mid n, m \geq 1\}$ 不是 RL。

(8) $\{xwx^T \mid x, w \in \Sigma^+\}$ 。

解: 该语言是 RL, 所以我们需要完整地给出 R_L 所确定的所有等价类。在给出这些等价类的过程中, 注意到该语言的句子的长度都至少为 3 是非常重要的。

正则语言的性质

第
5
章

R_L 将 $\{0, 1\}^+$ 分成如下 7 个等价类:

$$[\epsilon] = \{\epsilon\}$$

$$[0] = \{0\}$$

$$[1] = \{1\}$$

$$[01] = \{x \mid x \text{ 以 } 0 \text{ 开头, 以 } 1 \text{ 结尾的长度至少为 } 2\} \cup \{00\}$$

$$[10] = \{x \mid x \text{ 以 } 1 \text{ 开头, 以 } 0 \text{ 结尾的长度至少为 } 2\} \cup \{11\}$$

$$[00] = \{x \mid x \text{ 以 } 0 \text{ 开头, } 0 \text{ 结尾的长度至少为 } 3\}$$

$$[11] = \{x \mid x \text{ 以 } 1 \text{ 开头, } 1 \text{ 结尾的长度至少为 } 3\}$$

所以, R_L 的指数是有穷的, 由 Myhill-Nerode 定理, $\{xwx^T \mid x, w \in \Sigma^+\}$ 是 RL。

8. 判断下列命题, 并证明你的结论。

(1) RL 的每一个子集都是 RL。

解: 此结论是不对的。

显然, 对于任意一个字母表, 该字母表的克林闭包为 RL, 这个字母表上的其他任何语言都是其克林闭包的子集, 但是它们并不一定是 RL。

例如: $\{0, 1\}^*$ 是 RL, 但是它的子集 $\{0^n 1^n \mid n \geq 1\}$ 却不是 RL。

(2) 每一个 RL 都有一个正则的真子集。

解: 结论不正确。

该命题粗看起来好像是正确的: 有穷集合是 RL, 而语言是一个有穷集合或者是一个无穷集合, 每个集合都很容易找到一个有穷子集。但是, 其中存在这样一个问题, 并不是所有的有穷集合都有真子集。 就是这样的集合。

正则语言 不存在正则的真子集。

(3) $L = \{x \mid x = x^T\}$ 是 Σ 上的语言, L 是正则的。

解: 结论不正确。

这与第 1 题的第 (11) 小题基本上是一样的, 差别仅仅是这里的字母表并不是一个具体的字母表。不过, 取与 $z = (01)^N (10)^N$ 类似的串, 用泵引理很容易证明 L 不是 RL。另外, 根据这种特殊形式的串, 不难发现 R_L 将 Σ^* 划分成无穷多个等价类, 所以, 根据 Myhill-Nerode 定理, L 不是 RL。

9. 设 $L = \{x \mid x \text{ 中 } 0 \text{ 的个数不等于 } 1 \text{ 的个数}\}$ 是字母表 $\{0, 1\}$ 上的语言, 证明 L 不是 RL。

提示: 参考本章第 5 题的解法。

10. 证明: 无穷多个 RL 的并不一定是 RL。即 RL 对“无穷的并运算”是不封闭的。

提示: 一个无穷集合可以是由无穷多个有穷集合组成的。

$$\delta([q, p], a) = [\delta(q, a), \delta(p, a)]$$

16. 构造一个 DFA M , $L(M) = \{x000y \mid x, y \in \{0, 1\}^*\}$, 并且, M 中存在两个不同的状态 q 和 p , 满足 $[q] = [p]$ 。

解: 首先, 按照第 3 章给出的构造 DFA 的思路以及在那里所获得的经验, 读者应该比较容易地构造一个不一定满足要求的 DFA, 比如在原主教材的例 3-2 中, 我们曾经是按照如下方式确定状态的:

q_0 —— M 的启动状态。

q_1 —— M 读到了一个 0, 这个 0 可能是子串“000”的第 1 个 0。

q_2 —— M 在 q_1 后紧接着又读到了一个 0, 这个 0 可能是子串“000”的第 2 个 0。

q_3 —— M 在 q_2 后紧接着又读到了一个 0, 发现输入字符串含有子串“000”。因此, 这个状态应该是终止状态。

M 的状态转移函数:

$\delta(q_0, 0) = q_1$ —— M 读到了一个 0, 这个 0 可能是子串“000”的第 1 个 0。

$\delta(q_1, 0) = q_2$ —— M 又读到了一个 0, 这个 0 可能是子串“000”的第 2 个 0。

$\delta(q_2, 0) = q_3$ —— M 找到了子串“000”。

$\delta(q_0, 1) = q_0$ —— M 在 q_0 读到了一个 1, 它需要继续在 q_0 “等待”可能是子串“000”的第 1 个 0。

$\delta(q_1, 1) = q_0$ —— M 在刚刚读到了一个 0 后, 读到了一个 1, 表明在读入这个 1 之前所读入的 0 并不是子串“000”的第 1 个 0, 因此, M 需要重新回到状态 q_0 , 以寻找子串“000”的第 1 个 0。

$\delta(q_2, 1) = q_0$ —— M 在刚刚发现了 00 后, 读到了一个 1, 表明在读入这个 1 之前所读入的 00 并不是子串“000”的前两个 0, 因此, M 需要重新回到状态 q_0 , 以寻找子串“000”的第 1 个 0。

$\delta(q_3, 0) = q_3$ —— M 找到了子串“000”, 只用读入该串的剩余部分。

$\delta(q_3, 1) = q_3$ —— M 找到了子串“000”, 只用读入该串的剩余部分。

至此, 得到接受语言 $\{x000y \mid x, y \in \{0, 1\}^*\}$ 的 DFA: