2.1. (a) C语言的源字符集由96个字符组成:

①. 5个空白字符:空格' ',水平制表'\t',垂直制表'\v',换页'\f',换行'\n'
(在网上还搜到了回车'\r',但参考手册中并未提及)

② 10个数字字符:[0-9]

③. 52个字母: [A-Z] U [a-z]

④ 29个标点字符: _ { } [ ] # ( ) < > % : ; . ? * + - / ^ & | ~ ! = , \ " '

2.3 (d) 该正则表达式表示字母表{0,1}上由三个1和任意个(包括零个)0以任意顺序构成的所有串的集合.

2.4 解.(c) 依题.若X为没有一处相邻数字相同的所有数字串
则XX即为所求

从0开始归纳

$T_0 \to 0|\varepsilon$

$T_1 \to T_0?1(T_0 1)^* T_0? | T_0 | \varepsilon$

$T_2 \to T_1?2(T_1 2)^* T_1? | T_1 | \varepsilon$

......

$T_9 \to T_8?9(T_8 9)^* T_8? | T_8 | \varepsilon$

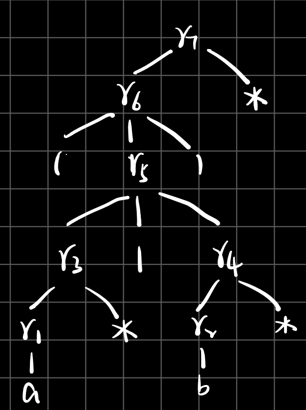
$Result \to T_9 T_9 | \varepsilon$

(g) 依题,该语言正规定义为

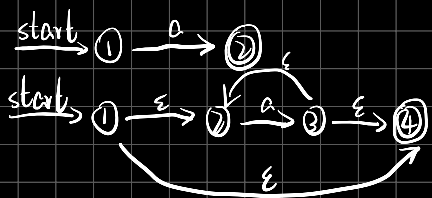
even_0_even_1 → $(00|11)^*((01|10)(00|11)^*(01|10)(00|11)^*)^*$

even_0_odd_1 → 1 even_0_even_1 | 0 $(00|11)^*$ (01|10) even_0_even_1

2.7 解.(b) 依题 $(a^*|b^*)^*$ 的分析树如右图

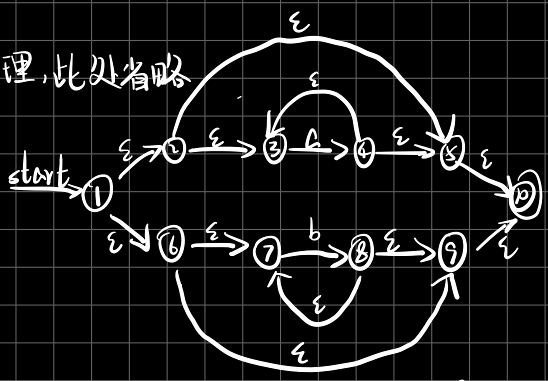


故 a 的NFA为

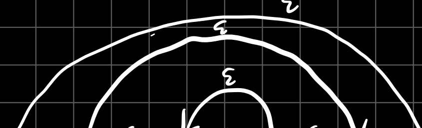


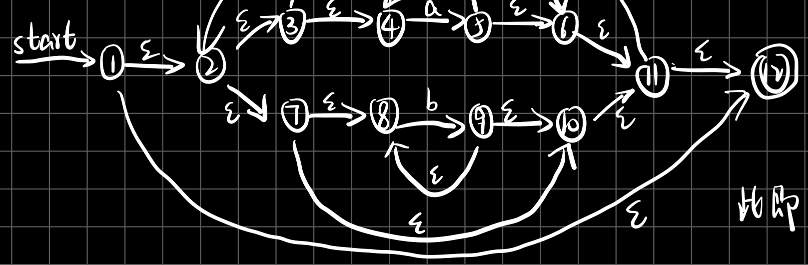
b, b^* 同理,此处省略

$a^*|b^*$:



$(a^*|b^*)^*$





此即 $(a^*|b^*)^*$ 的NFA

处理 ababbab 的转换序列为

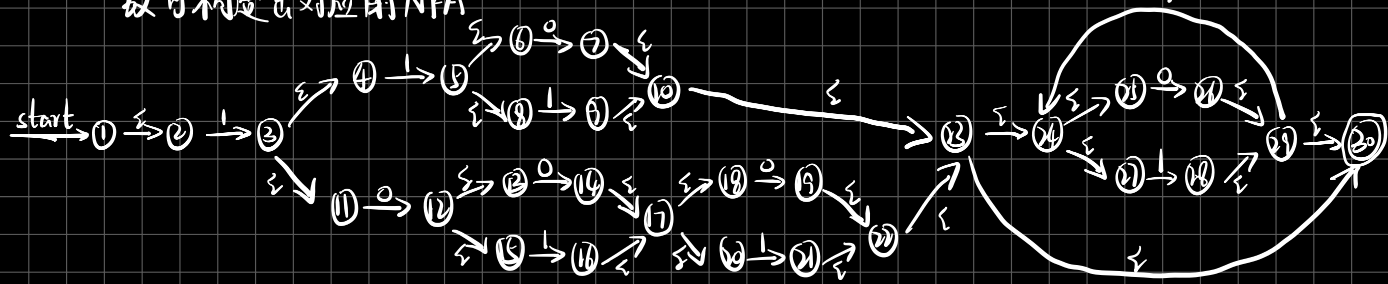
$\xrightarrow{start} 1 \xrightarrow{\varepsilon} 2 \xrightarrow{\varepsilon} \underbrace{3 \xrightarrow{\varepsilon} 4 \xrightarrow{a} 5 \xrightarrow{\varepsilon} 6 \xrightarrow{\varepsilon} 11}_{\text{设该段为X}} \xrightarrow{\varepsilon} 2 \xrightarrow{\varepsilon} \underbrace{7 \xrightarrow{\varepsilon} 8 \xrightarrow{b} 9 \xrightarrow{\varepsilon} 10 \xrightarrow{\varepsilon} 11}_{Y}$

$\xrightarrow{\varepsilon} 2 \xrightarrow{\varepsilon} X \xrightarrow{\varepsilon} 2 \xrightarrow{\varepsilon} Y \xrightarrow{\varepsilon} 2 \xrightarrow{\varepsilon} Y \xrightarrow{\varepsilon} 2 \xrightarrow{\varepsilon} X \xrightarrow{\varepsilon} 2 \xrightarrow{\varepsilon} Y \xrightarrow{\varepsilon} 12$

2.16 解. 依题, 可接受的输入为 110, 111, 1000, 1001, ……

其正则表达式为 $1\,(10|11|(00|01)(0|1))(0|1)^*$, 或化简为 $1\,(1(0|1)\,|\,0(0|1)(0|1))(0|1)^*$

故可构造出对应的NFA



故由子集构造算法

$A = \varepsilon\text{-closure}(\{1\}) = \{1,2\}$

$\varepsilon\text{-closure}(move(A,0)) = \phi$

$\varepsilon\text{-closure}(move(A,1)) = \varepsilon\text{-closure}(\{3\}) = \{3,4,11\} = B$

$\varepsilon\text{-closure}(move(B,1)) = \{5,6,8\} = C$

$\varepsilon\text{-closure}(move(B,0)) = \varepsilon\text{-closure}(\{12\}) = \{12,13,15\} = D$

$\varepsilon\text{-closure}(move(C,0)) = \{7,10,23,24,25,27,30\} = E_1$

$\varepsilon\text{-closure}(move(C,1)) = \{9,10,23,24,25,27,30\} = E_2$

$\varepsilon\text{-closure}(move(D,0)) = \{14,17,18,20\} = G_1$

$\varepsilon\text{-closure}(move(D,1)) = \{16,17,18,20\} = G_2$

$\varepsilon\text{-closure}(move(G_1,0)) = \{19,22,23,24,25,27,30\} = E_3$

$\varepsilon\text{-closure}(move(G_1,1)) = \{21,22,23,24,25,27,30\} = E_4$

$\varepsilon\text{-closure}(move(G_2,0)) = E_3$

$\varepsilon\text{-closure}(move(G_2,1)) = E_4$

对 $i = 1,2,3,4$ 有

$\varepsilon\text{-closure}(move(E_i,0)) = \{24,25,26,27,29,30\} = F_1$

$\varepsilon\text{-closure}(move(E_i,1)) = \{24,25,27,28,29,30\} = F_2$

对 $i=1,2$ 有

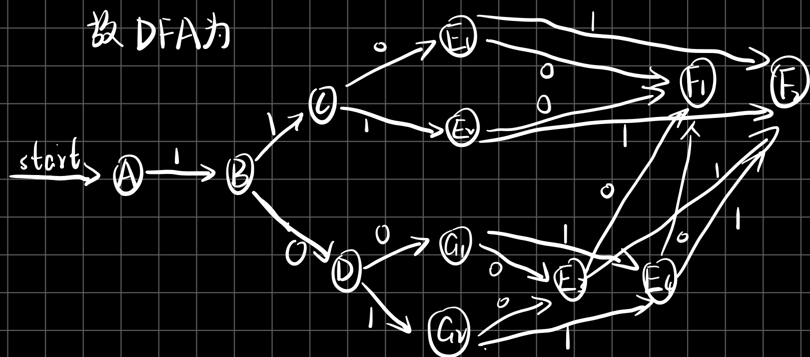
$$\varepsilon\text{-closure}(move(F_i,0)) = F_1$$

$$\varepsilon\text{-closure}(move(F_i,1)) = F_2$$

综上,可得DFA的转换表

state	input	
	0	1
A	ϕ	B
B	D	C
C	E_1	E_2
D	G_1	G_2
E_1	F_1	F_2
E_2	F_1	F_2
E_3	F_1	F_2
E_4	F_1	F_2
G_1	E_3	E_4
G_2	E_3	E_4
F_1	F_1	F_2
F_2	F_1	F_2

故DFA为



设 $S-F'=\{A,B,C,D,G_1,G_2\}$ $F'=\{E_1,E_2,E_3,E_4,F_1,F_2\}$

$\Pi=\{\{A,B,C,D,G_1,G_2\},\{E_1,E_2,E_3,E_4,F_1,F_2\}\}$

易知对 $\forall X\in F', y\in\{0,1\}, move(X,y)\subseteq F'$

而对于 $S-F'$,有 $\forall X\in\{C,G_1,G_2\}, y\in\{0,1\}$ $move(X,y)\subseteq F'$

$\forall X\in\{A,B,D\}, y\in\{0,1\}$ $move(X,y)\subseteq F'$

所以 $\Pi=\Pi_{new}=\{\{A,B,D\},\{C,G_1,G_2\},\{E_1,E_2,E_3,E_4,F_1,F_2\}\}$

而 $move(A,0), move(B,0)\in\{A,B,D\}$ $move(D,0)\notin\{A,B,D\}$

故 $\Pi=\Pi_{new}=\{\{A,B\},\{D\},\{C,G_1,G_2\},\{E_1,E_2,E_3,E_4,F_1,F_2\}\}$

$move(A,0)\subseteq\{A,B\}$, $move(B,0)\notin\{A,B\}$

故 $\Pi=\Pi_{new}=\{\{A\},\{B\},\{D\},\{C,G_1,G_2\},\{E_1,E_2,E_3,E_4,F_1,F_2\}\}$.

设 $1=\{A\}$, $2=\{B\}$, $3=\{D\}$

$4=\{C,G_1,G_2\}$, $5=\{E_1,E_2,E_3,E_4,F_1,F_2\}$

故最简DFA的转换表为

state	input	
	0	1
1	ϕ	2
2	3	4
3	4	4
4	5	5
5	5	5

所以DFA为