

编译原理和技术 B 作业 1

2.1. 从下列每种语言的参考手册确定它们形成输入字母表的字符集 (不包括那些只可以出现在字符串或注释中的字符).

(a) C

解: (a) 从 *C: A Reference Manual* 第五版第一部分第 2.1 节可得到形成 C 语言输入字母表的字符集包括

- 52 个大小写拉丁字母;
- 10 个阿拉伯数字;
- 空格;
- 水平制表符 HT, 垂直制表符 VT 和换页符 FF;
- 29 个特殊字符.

所以 C 语言输入字母表的字符集为 { a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z, A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, !, ", #, %, &, ', (,), *, +, ,, -, ., /, :, ;, <, =, >, ?, [\,], ^, _, {, |, }, ~, 空格, 水平制表符 HT, 垂直制表符 VT, 换页符 FF }

2.3. 叙述由下列正规式描述的语言.

(b) $((\varepsilon|0)1^*)^*$

解: (b) 因为 $\{\varepsilon, 0, 1\} \subset L((\varepsilon|0)1^*)$, 而显然 $L(((\varepsilon|0)1^*)^*) \subseteq L((0|1)^*)$, 所以有 $L(((\varepsilon|0)1^*)^*) = L((0|1)^*)$. 即 $((\varepsilon|0)1^*)^*$ 表示由 0 和 1 构成的所有串的集合, 包括空串.

2.4. 为下列语言写出正规定义.

(b) 按字典序排列的所有字母串.

(i) 字母表 $\{a, b\}$ 上, a 不会相邻出现的所有串.

解: (b) $answer \rightarrow A^*a^*B^*b^*\dots Z^*z^*$

(i) $answer \rightarrow a?(b|ba)^*$

2.7. 用算法 2.4 为下列正规式构造不确定有限自动机, 给出它们处理输入串 *ababbab* 的状态转换序列.

(c) $((\varepsilon|a)b^*)^*$

解: (c) 不确定有限自动机如下图 7.1, 状态转换序列为 $0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 6 \rightarrow 7 \rightarrow 8 \rightarrow 9 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 6 \rightarrow 7 \rightarrow 8 \rightarrow 7 \rightarrow 8 \rightarrow 9 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 6 \rightarrow 7 \rightarrow 8 \rightarrow 9 \rightarrow 10$.

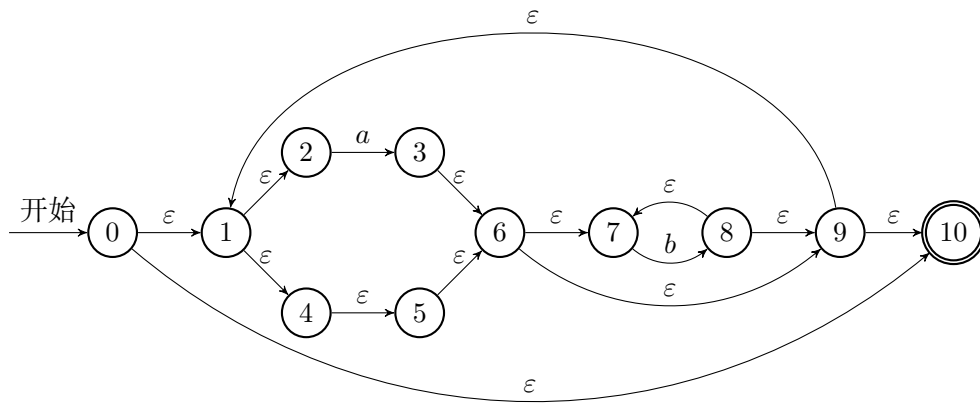


图 7.1: 识别 $((\varepsilon|a)b^*)^*$ 的 NFA

2.12. 为下列正规式构造最简的 DFA.

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

解: (a) 首先构造 NFA 如下图 12.1. 然后使用算法 2.2 将其转换为 DFA:

$$\varepsilon\text{-closure}(0) = \{0, 1, 2, 4, 7\} = A \quad (12.1)$$

$$\varepsilon\text{-closure}(\text{move}(A, a)) = \varepsilon\text{-closure}(\{3, 8\}) = \{1, 2, 3, 4, 6, 7, 8, 9, 11\} = B \quad (12.2)$$

$$\varepsilon\text{-closure}(\text{move}(A, b)) = \varepsilon\text{-closure}(\{5\}) = \{1, 2, 4, 5, 6, 7\} = C \quad (12.3)$$

$$\varepsilon\text{-closure}(\text{move}(B, a)) = \varepsilon\text{-closure}(\{3, 8, 10\}) = \{1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 13\} = D \quad (12.4)$$

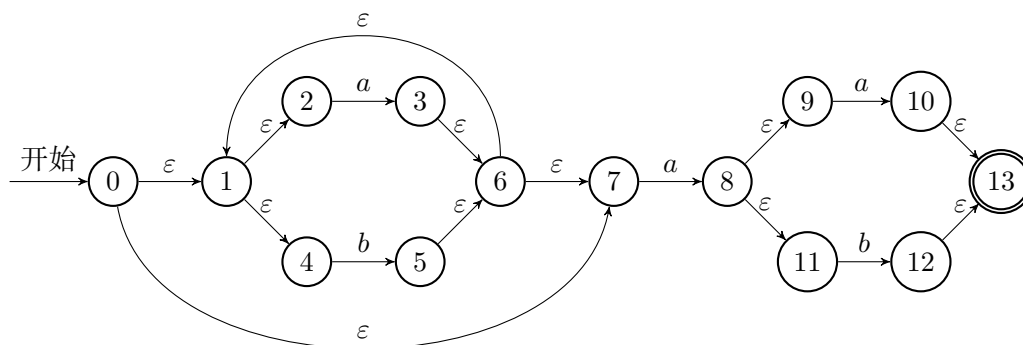
$$\varepsilon\text{-closure}(\text{move}(B, b)) = \varepsilon\text{-closure}(\{5, 12\}) = \{1, 2, 4, 5, 6, 7, 12, 13\} = E \quad (12.5)$$

$$\varepsilon\text{-closure}(\text{move}(C, a)) = \varepsilon\text{-closure}(\{3, 8\}) = B \quad (12.6)$$

$$\varepsilon\text{-closure}(\text{move}(C, b)) = \varepsilon\text{-closure}(\{5\}) = C \quad (12.7)$$

$$\varepsilon\text{-closure}(\text{move}(D, a)) = \varepsilon\text{-closure}(\{3, 8, 10\}) = D \quad (12.8)$$

$$\varepsilon\text{-closure}(\text{move}(D, b)) = \varepsilon\text{-closure}(\{5, 12\}) = E \quad (12.9)$$

图 12.1: 识别 $(a|b)^*a(a|b)$ 的 NFA

$$\varepsilon\text{-closure}(\text{move}(E, a)) = \varepsilon\text{-closure}(\{3, 8\}) = B \quad (12.10)$$

$$\varepsilon\text{-closure}(\text{move}(E, b)) = \varepsilon\text{-closure}(\{5\}) = C \quad (12.11)$$

A 是开始状态, D 和 E 是接受状态, 完整的转换表 $Dtran$ 如下表 12.1 所示, 构造出对应的 DFA 如图 12.2 所示.

状态	输入符号	
	a	b
A	B	C
B	D	E
C	B	C
D	D	E
E	B	C

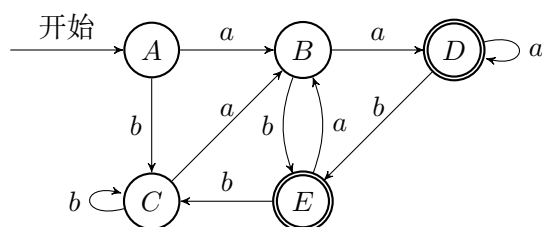
表 12.1: DFA 的转换表 $Dtran$ 

图 12.2: 子集构造法得到的 DFA

最后使用算法 2.3 化简得到的 DFA. 划分过程如下

$$\{A, B, C, D, E\} \begin{cases} \{A, B, C\} \\ \{D, E\} \end{cases} \begin{cases} \{A, C\} \\ \{B\} \\ \{D\} \\ \{E\} \end{cases} \quad (12.12)$$

选择 A 作为 $\{A, C\}$ 的代表, 其他状态不变, 可以得到最简自动机. 它的转换表如下表 12.2, 最简的 DFA 如下图所示 12.3.

状态	输入符号	
	a	b
A	B	A
B	D	E
D	D	E
E	B	A

表 12.2: DFA 的转换表 $Dtran$

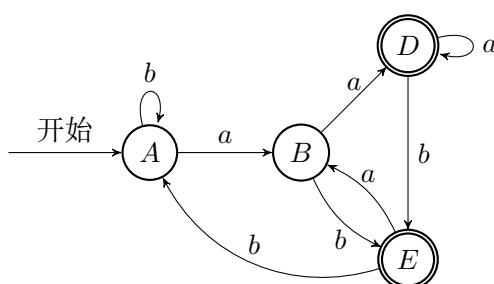


图 12.3: 最简的 DFA