

编译HW1

林宸昊 PB20000034.

2.1

(a)

1. 52个大小写拉丁字母

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

2. 10个数字

0 1 2 3 4 5 6 7 8 9

3. 空格

4. 水平制表符 (HT)、垂直制表符 (VT) 和换页符 (FF)

5. 29个特殊字符及其正式名称 (见表2-1)

源程序还要用某种方式分行, 为此可以使用字符、字符序列或源字符集之外的机制 (如记录结束符)。

表2-1 特殊字符

字 符	名 称	字 符	名 称	字 符	名 称
!	感叹号	+	加号	"	引号
#	数字号	=	等号	{	左花括号
%	百分号	-	波浪号	}	右花括号
^	折音符	[左方括号	,	逗号
&	和号]	右方括号	.	句号
*	星号	'	撇号	<	小于号
(左括号		竖线	>	大于号
_	下划线	\	反斜杠	/	除号
)	右括号	:	分号	?	问号
-	连字符	;	冒号		

* 根据《C语言参考手册第五版》, 输入字母字符集包括图中1、2、3、4、5以及如退格符、回车符的格式符。

* 其他特殊字符只能放在注释、字符型常量, 字符串型常量及文件名中。

2.3

(b) 字母表 $\{0, 1\}$ 上由0、1组成的所有字符串集合, 包括空串。

2.4

(b) 作如下定义。

letter-a \rightarrow A | a

letter-b \rightarrow B | b

letter-z \rightarrow z/z

最终目标定义则为

string \rightarrow (letter-a)* (letter-b)* ... (letter-z)*

(i)

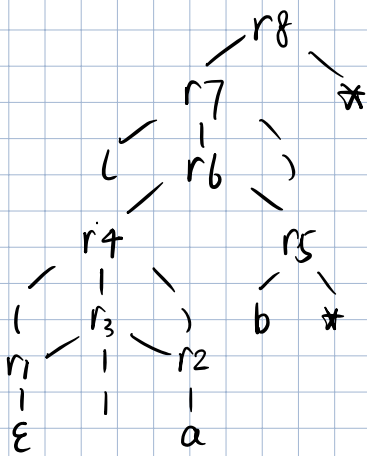
letter-a \rightarrow a

letter-b \rightarrow b

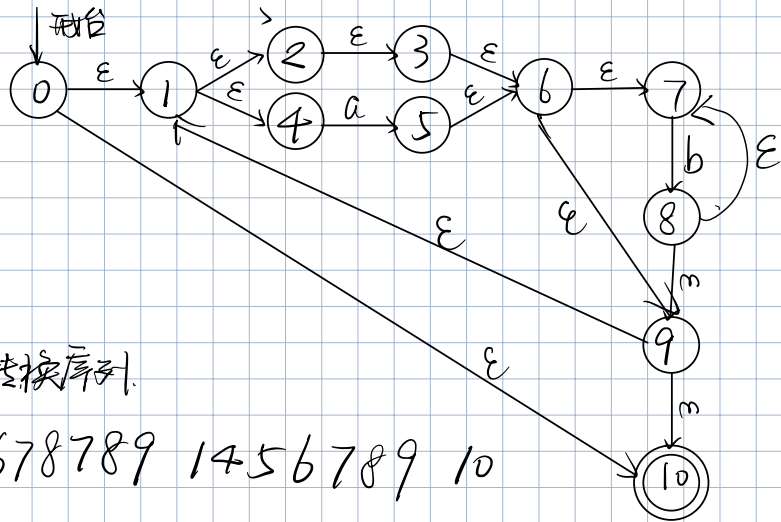
no-aa \rightarrow ((letter-a letter-b)* (letter-b)*)* (letter-a | ϵ)

2.7

(c) 对 $((\epsilon|a)b^*)^*$ 进行分解



NFA =

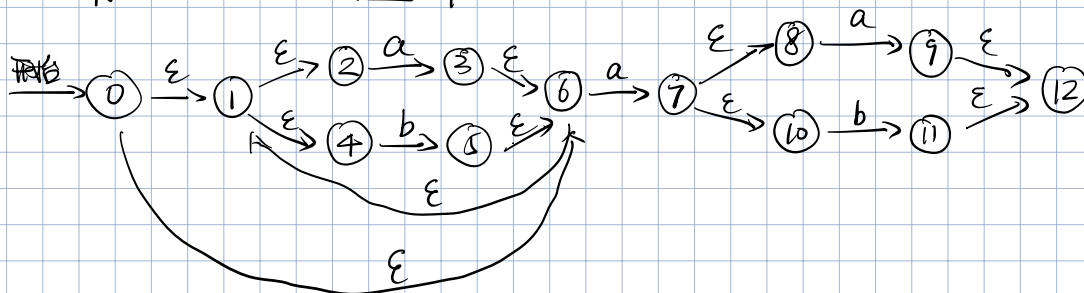


接收 ababbab 的状态转换序列

0 1 4 5 6 7 8 9 1 4 5 6 7 8 7 8 9 1 4 5 6 7 8 9 10

2.12

(a) 首先由正规式构造NFA



进行状态分类:

$$A = \{0, 1, 6, 2, 4\}$$

$$B = \{1, 2, 3, 4, 6, 7, 8, 10\} \text{ (move(A, a))}$$

$$C = \{1, 2, 4, 5, 6\} \text{ (move(A, b))}$$

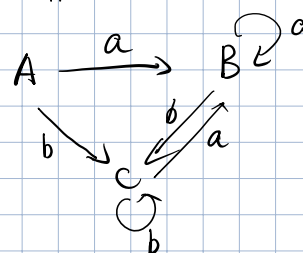
$$D = \{1, 2, 3, 4, 6, 7, 8, 9, 10, 12\} \text{ (move(B, a))}$$

$$E = \{1, 2, 4, 5, 6, 11, 12\} \text{ (move(B, b))}$$

得到转换表:

状态	a in	b in
A	B	C
B	D → B	E → C
C	B	C
D	D → B	E → C
E	B	C

则最终有:



易得 $B + D \rightarrow B$ $C + E \rightarrow C$