

# 第二次作业

2153067 王灏博

## 3.1.1

习题3.1.1 写出表示下列语言的正则表达式：

- \* a) 字母表 $\{a, b, c\}$ 上包含至少一个 $a$ 和至少一个 $b$ 的串的集合。
- b) 倒数第10个符号是1的0和1的串的集合。
- c) 至多只有一对连续1的0和1的串的集合。

a) 正则表达式为  $c^*(a(a+c)^*b + b(b+c)^*a)(a+b+c)^*$

b) 正则表达式为  $(0+1)^*1(0+1)^9$

c) 正则表达式为  $(10+0)^*(\epsilon + 1+11)(0+01)^*$

## 3.1.3

!! 习题3.1.3 写出表示下列语言的正则表达式：

- a) 不包含101作为子串的所有0和1的串的集合。
- b) 具有相同个数的0和1，使得在任何前缀中，0的个数不比1的个数多2，1的个数也不比0的个数多2，所有这种0和1的串的集合。
- c) 0的个数被5整除且1的个数是偶数的所有0和1的串的集合。

a) 正则表达式为  $\epsilon + (0+11^*00)^*(\epsilon + 11^* + 11^*0)$

b) 正则表达式为  $(01+10)^*$

c) 正则表达式为  $(00000 + 11 + (01+10)(11)^*10000 + (001 + (01+10)(11)^*(0+101))(11)^*(0+101))(11)^*1000 + (0001 + (01+10)(11)^*1001 + (001 + (01+10)(11)^*(0+101))(11)^*(0+101))(11)^*100 + (00001 + (01+10)(11)^*10001 + (001 + (01+10)(11)^*(0+101))(11)^*1001 + (0001 + (01+10)(11)^*1001 + (001 + (01+10)(11)^*(0+101))(11)^*(0+101))(11)^*(0+101))(11 + 00(11)^*10001 + 00(11)^*(0+101)(11)^*1001 + (00(11)^*1001 + 00(11)^*(0+101)(11)^*(0+101))(11)^*(0+101))(11)^*(0+101))^*(10+01+00(11)^*10000 + 00(11)^*(0+101)(11)^*1000 + (00(11)^*1001 + 00(11)^*(0+101)(11)^*(0+101))(11)^*100))^*$

### 3.1.4

！习题3.1.4 给出下列正则表达式语言的自然语言描述：

\* a)  $(1 + \epsilon)(00^*1)^*0^*$ 。

b)  $(0^*1^*)^*000(0 + 1)^*$ 。

c)  $(0 + 10)^*1^*$ 。

a)自然语言描述为：不包含相邻的1的0和1的串的集合

b)自然语言描述为：包含子串000的串的集合

c)自然语言描述为：要么没有连续的1的子串，要么存在一个连续的1的子串，并且这个连续的1的串后没有任何的0的串的集合

### 3.2.4

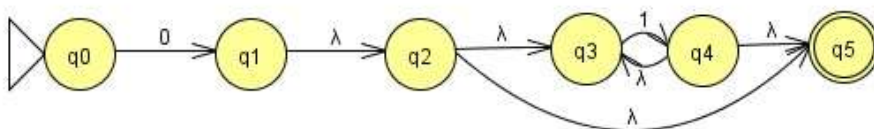
习题3.2.4 把下列正则表达式转化成带 $\epsilon$ 转移的NFA。

\* a)  $01^*$ 。

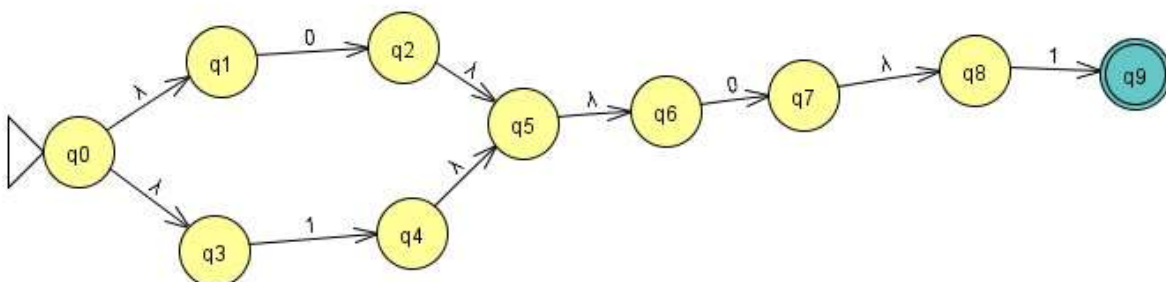
b)  $(0 + 1)01$ 。

c)  $00(0 + 1)^*$ 。

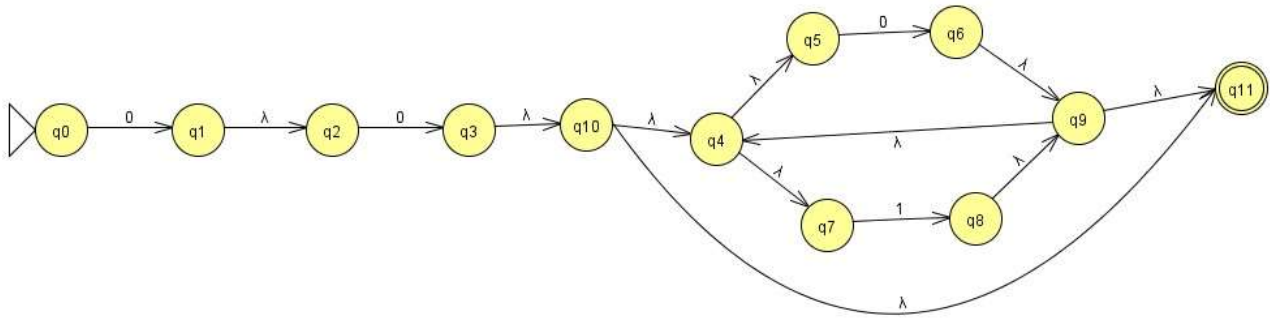
a)



b)



c)



### 3.2.6(1)

！习题3.2.6 设  $A = (Q, \Sigma, \delta, q_0, \{q_f\})$  是一个  $\epsilon$ -NFA，使得既没有进入  $q_0$  的转移，也没有离开  $q_f$  的转移。就  $L = L(A)$  而言，描述  $A$  的每一个下列修改所接受的语言：

- \* a) 通过增加从  $q_f$  到  $q_0$  的  $\epsilon$  转移，从  $A$  构造的自动机。
- \* b) 通过增加从  $q_0$  到每个从  $q_0$  可达（沿着标记包含  $\Sigma$  中符号和  $\epsilon$  的路径）的状态的  $\epsilon$  转移，从  $A$  构造的自动机。

a) 接受的语言是  $L^+$

b) 接受的语言是  $L$  的所有后缀的集合(包括  $L$  本身)

### 3.2.6(2)

- c) 通过增加从每个能沿着某条路径到达  $q_f$  的状态到  $q_f$  的  $\epsilon$  转移，从  $A$  构造的自动机。
- d) 通过同时做(b)和(c)的修改从  $A$  构造的自动机。

c) 接受的语言是  $L$  的所有前缀的集合(包括  $L$  本身)

d) 接受的语言是  $L$  的所有前缀和后缀的集合(包括  $L$  本身)

## 3.2.7

**!! 习题3.2.7** 把正则表达式转化为 $\epsilon$ -NFA的定理3.7的构造，有些地方可以化简。这里有三处：

1. 对于并运算符，不是构造新的初始状态和接受状态，而是把两个初始状态合并成一个具备两个初始状态的所有转移的状态。同样，合并两个接受状态，让所有的转移相应地进入合并状态。
2. 对于连接运算符，把第一个自动机与第二个自动机的接受状态合并。
3. 对于闭包运算符，只是增加从接受状态到初始状态的以及反方向的 $\epsilon$ 转移。

每一个这种化简本身仍然产生正确的构造；也就是说，对于任何正则表达式，得出的 $\epsilon$ -NFA接受这个表达式的语言。变化(1)、(2)和(3)的哪些子集可以一起用在构造中，对于每一个正则表达式，仍然产生正确的自动机？

变化(1),(2),(3)组成集合 $\{(1),(2),(3)\}$ 这个集合的子集

$\{\}$ ,  
 $\{(1)\}$ ,  
 $\{(2)\}$ ,  
 $\{(3)\}$ ,  
 $\{(1),(2)\}$ ,  
 $\{(1),(3)\}$

可以用于构造中而产生正确的 $\epsilon$ -NFA自动机

## 3.2.8

**\*!! 习题3.2.8** 给出一个算法：输入一个DFA  $A$ ，对于给定的 $n$ （与 $A$ 的状态个数无关），计算出 $A$ 所接受的长度为 $n$ 的串的个数。这个算法应当对于 $n$ 和 $A$ 的状态数来说都是多项式的。提示：使用定理3.4的构造所提示的技术。

令  $R_{ijm}^k$  表示从状态  $i$  到状态  $j$  长度为  $m$  的串的个数, 其中  $k$  表示经历的状态下标不大于  $k$ .  $i \leq j, m \leq n$   
下面使用归纳法来求解.

题目要求的问题即为求  $R_{ifn}^k$ , 其中  $k$  表示状态数,  $i$  表示起始状态  $f$  表示任一接受状态,  $n$  表示串的长度.

① 当  $k=0$  时,  $R_{iji}^0$  是从状态  $i$  到状态  $j$  的箭头的个数.

除了  $R_{iio}^0=1$  以外, 其余的  $R_{ijm}^0=0$ .

②  $R_{ijm}^k$  是所有  $R_{ijm}^{k-1}$  的和

其中, 某一个  $R_{ijm}^{k-1}$  可以表示为

$$R_{ikp_1}^{k-1} \times R_{k_1k_2}^{k-1} \times \dots \times R_{k_{p-1}k_p}^{k-1} \times R_{k_pj}^{k-1}$$

( $\sum_{i=1}^p p_i = m$ )

则答案即为所有  $R_{ifn}^k$  的和,  $f$  表示任一接受状态,  $n$  表示串的长度

## 3.3.1和3.3.2

**！习题3.3.1** 给出一个正则表达式, 来描述所能想到的所有不同形式的电话号码。考虑国际号码以及不同国家有不同位数的区号和本地电话号码。

**！！习题3.3.2** 给出一个正则表达式, 来表示在招聘广告中可能出现的薪水。考虑可能按小时、周、月或年发放的薪水。这些薪水可能有也可能没有\$ (如美元) 符号或其他单位 (如后面跟着的“K”)。可能有一个或多个邻近的单词标志着薪水。提示: 查看报纸上的分类广告或在线职位列表, 来获得一些关于什么样的模式可能有用的想法。

### 3.3.1

根据资料显示合法的区号由最多三位数字组成, 其中第一位是除了0以外的九个数字的任意一个合法的总长度(不包括'+')为[8,15]中的一个整数

设  $\sum_a = \{0,1,2,3,4,5,6,7,8,9\}$

设  $\sum_b = \{1,2,3,4,5,6,7,8,9\}$

所以匹配所有国家手机号码的正则表达式为:

$$'+' (\sum_b) ((\sum_a)^7 + (\sum_a)^8 + (\sum_a)^9 + (\sum_a)^{10} + (\sum_a)^{11} + (\sum_a)^{12} + (\sum_a)^{13} + (\sum_a)^{14})$$

### 3.3.2

从一些招聘网站上获得的招聘信息如图所示

 Up to \$2,300 a month

- Office and student administration support i) Flags maintenance (e.g. National and School Flags); ii) Procurement and management of stationery and office...

### \$1,800 - \$2,800 a month

- Responsible for order processing of print & digital **newspaper**, school and magazine titles.
- Work closely with different stakeholders to ensure all orders are...

### From \$11 an hour

- Engage in recreational activities enjoyed by clients e.g. reading the **newspaper**, playing musical instruments, singing, playing simple card or board games etc.

根据这些信息，尝试构造一个正则表达式来尽可能涵盖多的薪资：

设  $\sum_a = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

设  $\sum_b = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$

$(\epsilon + 'Up\ to')(\epsilon + '$)((\sum_b((\sum_a)^3 + (\sum_a)^4 + (\sum_a)^5)' - ' \sum_b((\sum_a)^3 + (\sum_a)^4 + (\sum_a)^5)'a\ month') + ('From' \sum_b \sum_a)'an\ hour')$