

2019 级弘毅班《编译原理》第一次练习

注意：2022 年 03 月 21 日上课时交作业！

一、为下述语言集合设计 DFA：

- (1) 所有的以 a, b 和 c 组成的字符串，其中， a, b 和 c 均不连续出现。
- (2) 所有的以 a 和 b 组成的字符串，以 $abbab$ 结尾。
- (3) 所有的以 a 和 b 组成的字符串，其中 a 出现的次数除 3 后的余数为 2。
- (4) 所有的以 a 和 b 组成的字符串，其中没有 aba 子串。
- (5) 所有的以 a 和 b 组成的字符串，其中至少有一个 aba 子串。
- (6) 所有的以 a 和 b 组成的字符串，其中子串 ab 出现的次数为偶数。

二、为题一所有的语言集合编写正规表达式。试用 JFLAP 验证你的设计，即其等价的最小 DFA 应与题一一致。请用 JFLAP(java -jar JFLAP.jar) 工具检验你的设计。

注意：JFLAP 输入正规表达式时，并运算用 “+” 表示，空串用 “!” 表示。

三、现给下述正规表达式：

$$(ab|b)^*aa(ba|b)^*$$

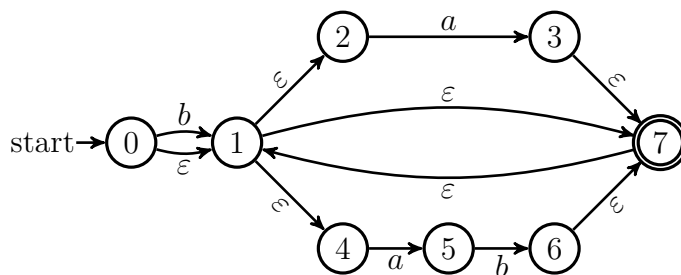
试对上述正规表达式：

- (1) 用 Thompson 构造法将上述正规表达式转化为 NFA；
- (2) 用子集构造法将上述的 NFA 转化为 DFA；
- (3) 用状态的划分法将上述的 DFA 的状态最小化；
- (4) 用自然语言描述正规表达式所表达的语言。

请用 JFLAP(java -jar JFLAP.jar) 工具检验你的设计。

四、设 NFA N 的状态转换图如下所示：

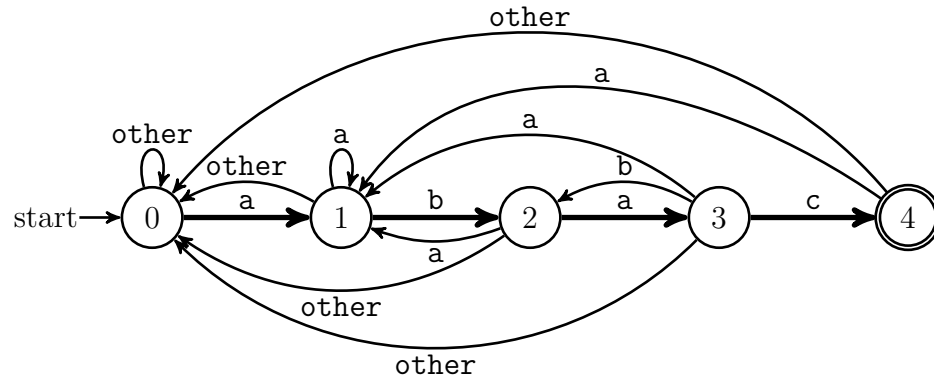
(2021 年考试试题)



- (1) 试写出 NFA N 接受字符串 “baab” 的过程；
- (2) 设用子集构造法求出的与 NFA N 等价的 DFA M 有 4 个状态 A, B, C 和 D ，其中 $A = \varepsilon\text{-closure}(\{0\})$ ，状态转换函数 $Dtrans(A, a) = B$ ， $Dtrans(A, b) = C$ 试求与状态 A, B, C 和 D 所对应的 NFA N 的状态集，并画出 DFA M 的状态转换图；
- (3) 求 DFA M 的最小状态自动机；

- (4) 试用自然语言描述 NFA N 所接受的语言;
- (5) 求正则表达式 r , 使得 $L(r) = L(N)$.

五、KMP 算法将搜索关键字的时间转换为线性时间, 设 t 是要搜索的关键字, 该算法就是一个接受以 t 为后缀的自动机。现需要设计一个函数 `int match(char *s)`, 函数对字符串 s 扫描, 如果发现有子串 “abac”, 则返回 0, 否则返回 1, 对应的接受 abac 为后缀的自动机状态图如下所示:



- (1) 将上述状态图转换为流程图;
- (2) 利用 C 语言的控制流结构实现函数 `int match(char *s);`
- (3) 思考: 为什么一般的正规表达式匹配不能用 KMP 算法。