

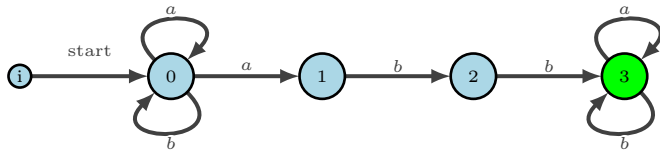
编译原理 HW2

肖桐 PB18000037

2020 年 9 月 26 日

NFA:

由 $(a|b)^*$ 可以知道, 处于开始状态时, 遇到任意个 a, b 均可继续保持在开始状态, 因此开始状态需要两个指向自己的箭头. 同时因为 abb 的存在, 开始状态还需要能够接受字符流 abb 之后进入接收状态. 进入接收状态之后, 不论再接收到多少个 a, b 都仍然处在接收状态. 即可得到以下NFA: (图中用绿色顶点表示接收状态, 同时因为要显示出 start 的箭头, 不得不加上最开始的顶点 i, 可忽略)

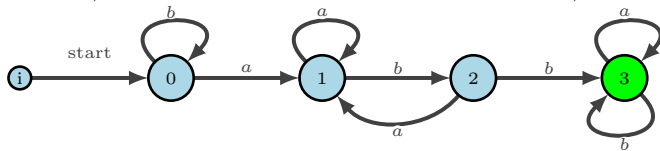


DFA:

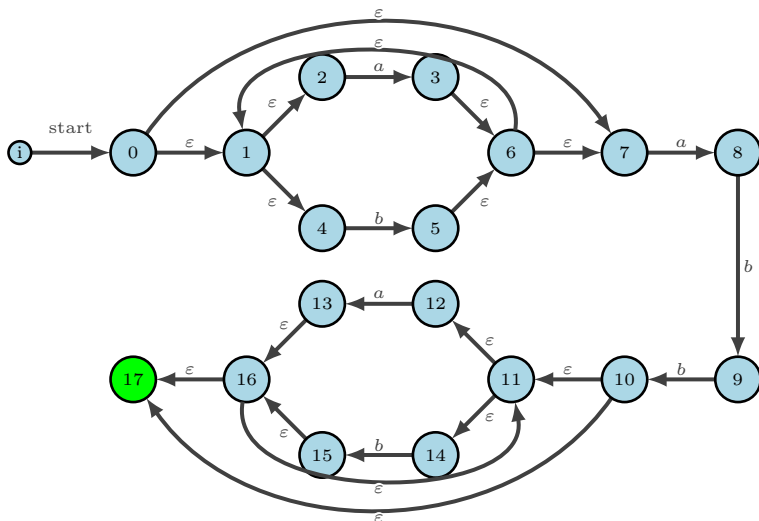
因为DFA要求每个状态经一个字符后只能到达一个状态, 因此在原NFA中, 状态0需要减少一个带 a 的箭头. 显然经 a 到达状态1的箭头不能减去, 否则无法到达接收状态, 因此删去指向状态0自身的带 a 箭头.

同时状态1缺少一个带 a 的箭头, 对正则表达式分析知, 接受任意个 a (至少有一个)与接受1个字符 a 对于正则表达式而言是等价的, 因此从状态1引出的带 a 的箭头应指向1自身.

同样的, 状态2也缺少一个带 a 的箭头, 对正则表达式分析知, 再识别字符串 ab 之后若再遇到字符 a , 则字符串 ab 只能算是满足正则表达式 $(a|b)^*$ 的字符串, 而不是满足正则表达式 abb 的字符串, 因此此时与仅识别了字符串 a 是等价的, 因此从2引出的带 a 箭头应当指向状态1, 故可得到以下的DFA:



按算法, 由正则表达式 $(a|b)^*abb(a|b)^*$ 得到的NFA:



下面用子集构造法求出该NFA对应的DFA(不一定最简).

记初始状态0的 ε -闭包为 $A = \{0, 1, 2, 4, 7\}$.

若集合 T 为从状态集合 A 中的某一状态 s 出发, 通过字符 a, ε 转换可以到达的NFA状态组成的集合, 则记 $A \xrightarrow{a} T$.

$$A \xrightarrow{a} \{1, 2, 3, 4, 6, 7, 8\} = B$$

$$A \xrightarrow{b} \{1, 2, 4, 5, 6, 7\} = C$$

$$B \xrightarrow{a} \{1, 2, 3, 4, 6, 7, 8\} = B$$

$$B \xrightarrow{b} \{1, 2, 4, 5, 6, 7, 9\} = D$$

$$C \xrightarrow{a} \{1, 2, 3, 4, 6, 7, 8\} = B$$

$$C \xrightarrow{b} \{1, 2, 4, 5, 6, 7\} = C$$

$$D \xrightarrow{a} \{1, 2, 3, 4, 6, 7, 8\} = B$$

$$D \xrightarrow{b} \{1, 2, 4, 5, 6, 7, 10, 11, 12, 17\} = E$$

$$E \xrightarrow{a} \{1, 2, 3, 4, 6, 7, 8, 11, 12, 13, 14, 16, 17\} = F$$

$$E \xrightarrow{b} \{1, 2, 4, 5, 6, 7, 11, 12, 14, 15, 16, 17\} = G$$

$$F \xrightarrow{a} \{1, 2, 3, 4, 6, 7, 8, 11, 12, 13, 14, 16, 17\} = F$$

$$F \xrightarrow{b} \{1, 2, 4, 5, 6, 7, 9, 11, 12, 14, 15, 16, 17\} = H$$

$$G \xrightarrow{a} \{1, 2, 3, 4, 6, 7, 8, 11, 12, 13, 14, 16, 17\} = F$$

$$G \xrightarrow{b} \{1, 2, 4, 5, 6, 7, 11, 12, 14, 15, 16, 17\} = G$$

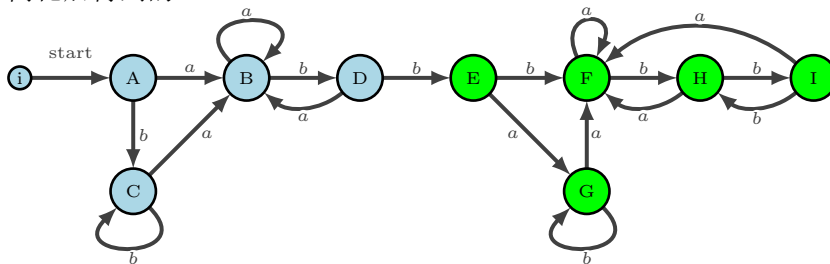
$$H \xrightarrow{a} \{1, 2, 3, 4, 6, 7, 8, 11, 12, 13, 14, 16, 17\} = F$$

$$H \xrightarrow{b} \{1, 2, 4, 5, 6, 7, 10, 11, 12, 14, 16, 17\} = I$$

$$I \xrightarrow{a} \{1, 2, 3, 4, 6, 7, 8, 11, 12, 13, 14, 16, 17\} = F$$

$$I \xrightarrow{b} \{1, 2, 4, 5, 6, 7, 9, 11, 12, 14, 15, 16, 17\} = H$$

简化后得到的DFA:



初始时, 只分可接收状态组和不可接收状态组.

因为从状态 D 出发经字符 b 可到达可接收状态, 而从状态 A, B, C 出发经任意字符都不能达到可接收状态, 因此 D 与 A, B, C 都是可分辨的, 因此 D 单独成为一组状态.

同理: 从状态 B 出发可以到达状态 D (另一个组), 而从 A, C 出发经任意字符都不能达到 D , 因此 B 与 A, C 也是可分辨的, 故 B 也单独成为一组状态.

最后, 因为 A, C 经字符 a 均到达 B 所在的状态组(就是 B 本身), 经字符 b 均到达 A, C 所在的状态组(就是 A, C 本身), 因此 A, C 是不可分辨的, 将 A, C 进行合并, 合并为 A .

同理: 可接收状态 E, F, G, H, I 之间也是不可分辨的, 故将这5个状态合并为 E . 最终可得到最简DFA:

