

作业 L2.1 - L2.3：正则文法与语法范畴

2025 年 12 月 16 日

题目 1

构造一个 DFA 接受如下文法定义的语言：

$$S \rightarrow abA, \quad A \rightarrow baB, \quad B \rightarrow aA \mid bb.$$

Solution.

该右线性文法生成的串必以 $S \Rightarrow abA \Rightarrow abbaB$ 开头，在 B 处要么选择 bb 结束，要么选择 aA 并再次经过 $A \Rightarrow baB$ 回到 B 。因此语言可写为：

$$L = abba(ab\alpha)^*bb.$$

据此构造 DFA（含陷阱态 \perp ）。令状态表示“已匹配到的固定前缀/循环位置”：

$$q_0 \xrightarrow{a} q_1 \xrightarrow{b} q_2 \xrightarrow{b} q_3 \xrightarrow{a} q_4$$

其中 q_4 表示已经读到 $abba$ 且当前位于循环入口（对应非终结符 B 的位置）。循环块 aba 由 $q_4 \xrightarrow{a} q_5 \xrightarrow{b} q_6 \xrightarrow{a} q_4$ 实现；终止块 bb 由 $q_4 \xrightarrow{b} q_7 \xrightarrow{b} q_8$ 实现，其中 q_8 为唯一接受态。所有未定义的转移均进入陷阱态 \perp （并在 \perp 上自环）。

完整转移（字母表 $\{a, b\}$ ）：

	a	b
q_0	q_1	\perp
q_1	\perp	q_2
q_2	\perp	q_3
q_3	q_4	\perp
q_4	q_5	q_7
q_5	\perp	q_6
q_6	q_4	\perp
q_7	\perp	q_8
q_8	\perp	\perp
\perp	\perp	\perp

该 DFA 恰好接受 $abba(ab\alpha)^*bb$ 。

题目 2

构造产生语言 $L(aa^*(ab+a)^*)$ 的正则文法。

Solution.

正则表达式 $aa^*(ab+a)^*$ 等价于 $a^+(ab|a)^*$ (至少一个 a, 之后由若干个 a 或 ab 拼接)。给出一个右线性文法 (正则文法):

$$\begin{aligned} S &\rightarrow aS \mid aR, \\ R &\rightarrow aR \mid aB \mid \epsilon, \\ B &\rightarrow bR. \end{aligned}$$

其中 S 保证至少产生一个 a; 在 R 中可以选择产生 a (用 $R \rightarrow aR$) 或产生 ab (用 $R \rightarrow aB \rightarrow abR$), 并可用 $R \rightarrow \epsilon$ 结束。

题目 3

构造产生语言 $L((aab^*ab)^*)$ 的右线性文法。

Solution.

一个块为 $aa b^* ab$, 整体是其 Kleene 星, 可产生空串。给出右线性文法:

$$\begin{aligned} S &\rightarrow \epsilon \mid aA, \\ A &\rightarrow aB, \\ B &\rightarrow bB \mid aC, \\ C &\rightarrow bS. \end{aligned}$$

推导: $S \Rightarrow aA \Rightarrow aaB \Rightarrow aab^*aC \Rightarrow aab^*abS$, 回到 S 继续重复或用 ϵ 结束。

题目 4

设文法 G 的产生式如下, 给出 G 的每个语法范畴 (非终结符) 代表的集合:

$$\begin{aligned} S &\rightarrow aSa \mid aaSaa \mid aAa \\ A &\rightarrow bA \mid bbbA \mid bB \\ B &\rightarrow cB \mid cC \\ C &\rightarrow ccC \mid DD \\ D &\rightarrow dD \mid d \end{aligned}$$

Solution.

(1) D :

$$D \rightarrow dD \mid d \Rightarrow L(D) = \{d^n \mid n \geq 1\}.$$

(2) C :

$$C \rightarrow ccC \mid DD.$$

其中 DD 连接两个 d^+ , 因此 $L(DD) = \{d^n \mid n \geq 2\}$ 。再在前面重复添加 cc , 得

$$L(C) = \{(cc)^k d^n \mid k \geq 0, n \geq 2\} = \{c^{2k} d^n \mid k \geq 0, n \geq 2\}.$$

(3) B :

$$B \rightarrow cB \mid cC \Rightarrow L(B) = \{c^m x \mid m \geq 1, x \in L(C)\}.$$

因此

$$L(B) = \{c^t d^n \mid t \geq 1, n \geq 2\}.$$

(4) A :

$$A \rightarrow bA \mid bbbA \mid bB \Rightarrow A \Rightarrow (b \mid bbb)^* bB.$$

即先产生至少一个 b , 再进入 B , 于是

$$L(A) = \{b^m c^t d^n \mid m \geq 1, t \geq 1, n \geq 2\}.$$

(5) S :

$$S \rightarrow aSa \mid aaSaa \mid aAa$$

表示在两端对称包裹 1 个或 2 个 a , 最终落到 aAa 。因此

$$L(S) = \{a^k x a^k \mid k \geq 1, x \in L(A)\}.$$

代入 $L(A)$ 得

$$L(S) = \{a^k b^m c^t d^n a^k \mid k \geq 1, m \geq 1, t \geq 1, n \geq 2\}.$$