

练习 4.1.1: 考虑上下文无关文法 $S \rightarrow +SS \mid *SS \mid a$ 以及串 $*+*aaaa$

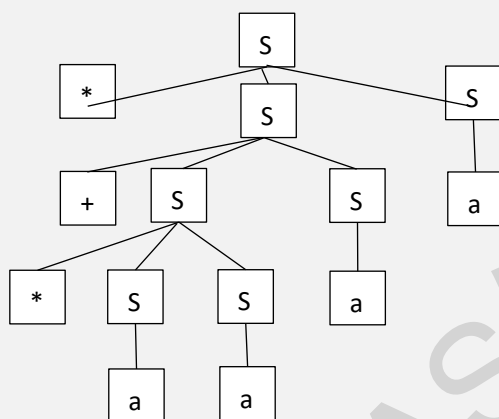
- 1) 给出这个串的一个最左推导
- 2) 给出这个串的一个最右推导
- 3) 给出这个串的一颗语法分析树
- 4) 这个文法是否是二义性的? 证明你的回答 (选作)
- 5) 这个文法生成的语言是什么?

答案

1) $S \Rightarrow *SS \Rightarrow *+SSS \Rightarrow *+*SSSS \Rightarrow *+*aSSS \Rightarrow *+*aaSS \Rightarrow *+*aaaS \Rightarrow *+*aaaa$

2) $S \Rightarrow *SS \Rightarrow *Sa \Rightarrow *+SSa \Rightarrow *+Saa \Rightarrow *+*SSaa \Rightarrow *+*Saaa \Rightarrow *+*aaaa$

3)



4) 不具有二义性。证明:

① 先证明一个该文法产生串的长度的结论: 设串的推导过程中使用产生式 $S \Rightarrow +SS$ 和 $S \Rightarrow *SS$ 的次数为 m , 则串的长度 $L=2*m+1$, 且串中包含 m 个运算符(+或*)和 $(m+1)$ 个 a 。

(1) 当 $m=0$ 时, 仅有 $S \Rightarrow a$ 一种情况, $L=1$, 串由 1 个 a 和 0 个运算符构成, 结论成立;

(2) 设当 $m < k(k \geq 1)$ 时结论成立, 则当 $m=k$ 时, 第一步推导必然为

$$S \Rightarrow S_1 S_2 op$$

op 为+或*, S 下标仅用于区分 S 的多次出现。设 $S_1 \xRightarrow{*} \alpha$, $S_2 \xRightarrow{*} \beta$, α 、 β 均为使用 $S \Rightarrow S_1 S_2 op$ 少于 k 次得到的串, 设二者推导过程中分别使用该产生式 k_1 和 k_2 次, 根据假设:

$$L(\alpha) = 2*k_1 + 1, L(\beta) = 2*k_2 + 1$$

串长度 $L = L(\alpha) + L(\beta) + 1 = 2*(k_1 + k_2 + 1) + 1 = 2*k + 1$; 且串中 a 的个数为 $(k_1 + 1) + (k_2 + 1) = k + 1$; 运算符个数为 $k_1 + k_2 + 1 = k$, 故结论成立。

② 下面证明该文法无二义性, 对串的长度做归纳。由前述证明可知, 该文法产生的串长 L 可为任意非负奇数。对由该文法得到的长度为 $L=2*k+1$ 串 ω :

(1) 当 $k=0$ 时, $L=2*0+1=1$, 只有 $S \Rightarrow a$ 一种情况, 显然没有二义性。

(2) 设当 $k < n$ 时结论成立。 $S \xRightarrow{*} \omega$, 根据 ω 首位运算符可确定第一步推导使用的产生式, 不妨设为:

$$S \Rightarrow +S_1 S_2$$

从前向后处理串 ω , 除去首位的运算符, 找到可以由 S 推导出的最短的串 α , 设 α 长

度为 m_1 ，由前述结论可知 $m_1=2*k_1+1$ ，且 α 包含 k_1 个运算符与 (k_1+1) 个 a ，由归纳假设可知 α 无二义性，存在唯一的最左推导 $S \xRightarrow[lm]{*} \alpha$ ；

设串 ω 剩余部分为 β ，设 β 长度为 m_2 ，同理可知 $m_2=2*k_2+1$ ， β 包含 k_2 个运算符与 (k_2+1) 个 a ，存在唯一最左推导 $S \xRightarrow[lm]{*} \beta$ ，且满足 $k=k_1+k_2$ 。

此时串 ω 可表示成如下形式：

$$\omega = +\alpha\beta$$

故存在唯一的最左推导：

$$S \xRightarrow[lm]{*} +SS \xRightarrow[lm]{*} +\alpha S \xRightarrow[lm]{*} +\alpha\beta$$

此时仍不存在二义性。

综上所述，该文法不具有二义性。

□

5) 由字符 a 与运算符 $+$ 、 $*$ 构成的前缀表达式。

练习 4.1.2：设计文法表示下面的串组成的语言：由二元运算符 \oplus 、 \otimes 和变量 x 组成的表达式，运算符右结合且 \oplus 的优先级高于 \otimes

$$\begin{aligned} S &\rightarrow T | T \otimes S \\ T &\rightarrow x | x \oplus T \end{aligned}$$

练习 4.1.3：下面是一个只包含符号 a 和 b 的正则表达式的文法，其中用 $+$ 替代表示并运算的字符 $|$ ，以避免和文法中作为元符号使用竖线混淆

$$\begin{aligned} \text{rexpr} &\rightarrow \text{rexpr} + \text{rterm} \mid \text{rterm} \\ \text{rterm} &\rightarrow \text{rterm} \text{rfactor} \mid \text{rfactor} \\ \text{rfactor} &\rightarrow \text{rfactor} * \mid \text{rprimary} \\ \text{rprimary} &\rightarrow a \mid b \end{aligned}$$

- 1) 对该文法提取左公因子
- 2) 提取左公因子的变换能使这个文法适用于自顶向下的语法分析技术吗？
- 3) 将提取了左公因子的文法继续消除左递归
- 4) 此时得到的文法适用于自顶向下的语法分析吗

- 1、没有左共因子
- 2、不能，因为文法还存在左递归
- 3、

$$\begin{aligned} \text{rexpr} &\rightarrow \text{rterm} \text{rexpr}' \\ \text{rexpr}' &\rightarrow +\text{rterm} \text{rexpr}' \mid \varepsilon \\ \text{rterm} &\rightarrow \text{rfactor} \text{rterm}' \\ \text{rterm}' &\rightarrow \text{rfactor} \text{rterm}' \mid \varepsilon \\ \text{rfactor} &\rightarrow \text{rprimary} \text{rfactor}' \\ \text{rfactor}' &\rightarrow * \text{rfactor}' \mid \varepsilon \\ \text{rprimary} &\rightarrow a \mid b \end{aligned}$$

- 4、合适