

## 2019 级弘毅班《编译原理》第二次练习

2022 年 04 月 18 日交作业.

一、 设有下述四组文法和其对应的终结符号串:

- (a)  $S \rightarrow 0S1 \mid 01$  with string 000111.
- (b)  $S \rightarrow +SS \mid *SS \mid a$ , string:  $+*aaa$ ;
- (c)  $S \rightarrow S(S)S \mid \varepsilon$ , string:  $((\ )(\ ))$ ;
- (d)  $S \rightarrow S+S \mid SS \mid (S) \mid S* \mid a$ , string:  $(a+a)*a$ .

试对上述文法和相应的字符串:

- (1) 给出一个最左推导;
- (2) 给出一个最右推导;
- (3) 画出其对应的语法树;
- (4) 对应文法是否是二义文法, 为什么;
- (5) 描述对应文法所生成的语言;
- (6) 消除对应文法的左递归和左公因子.

二、 试为下列语言设计上下文无关文法并判断该语言是否为正则语言;

- (1) 所有的有  $a$  和  $b$  组成的字符串, 其中, 每个  $b$  之前一定至少有一个  $a$ ;
- (2) 所有的有  $a$  和  $b$  组成的字符串, 其中  $a$  出现的次数比  $b$  出现的次数多;
- (3)  $\{a^m b^n c^p d^q \mid m, n, p, q \in \mathbb{N} \wedge m+n = p+q\}$ ; (提示: 分情况考虑:  $1/ m \leq q: S \rightarrow aSd \mid B$  with  $B = \{b^n c^p d^q \mid n = p+q\}$ . 这样  $B \rightarrow bBd \mid C$  with  $C = \{b^n c^p \mid n = p\}$ ;  $2/ m > q: \dots$ )
- (4)  $\{a^m b^n \mid m, n \in \mathbb{N} \wedge 2m = 3n + 1\}$  (提示: 需要找出  $a$  增长与  $b$  增长之关系, 设  $a$  增长  $p$  个,  $b$  增长  $q$  个, 则  $2(m+p) = 3(n+q) + 1$ , 这样  $2p = 3q$ );
- (5) 所有的有  $a$  和  $b$  组成的字符串, 其中  $a$  和  $b$  出现的次数不等.
- (6) 所有的以  $a$ 、 $b$  和  $c$  组成的字符串, 其中, 没有  $abb$  子串;
- (7)  $\{a^m b^n \mid m > n \wedge m-n \text{ 是偶数}\}$ .

三、 文法符号  $X$  (终结符或非终结符) 是无效的当且仅当不存在  $S \xRightarrow{*} wXy \xRightarrow{*} wxy$ . 即  $X$  不可能出现在推出语句的推导序列中.

- (1) 给出一个找出文法所有无效符号的算法.  
提示: 定义两个集合:

$$\mathcal{E} = \{U \mid U \in N \wedge U \xRightarrow{*} w \in T^*\}$$

$$\mathcal{A} = \{U \mid U \in N \wedge S \xRightarrow{*} \alpha U \beta\}$$

即  $\mathcal{E}$  是可终结非终结符;  $\mathcal{A}$  是可达非终结符. 利用类似 First 集和 Follow 集的算法求出  $\mathcal{E}$  和  $\mathcal{A}$ .

- (2) 用下述文法检验你的算法:

$$S \rightarrow 0 \mid A$$

$$A \rightarrow AB$$

$$B \rightarrow 1$$

四、对下述文法消除左递归和左公因子、求 First 和 Follow 集、构造 LL(1) 分析表

- (1)  $S \rightarrow SS+ \mid SS* \mid a.$
- (2)  $S \rightarrow 0S1 \mid 01.$
- (3)  $S \rightarrow S(S)S \mid \varepsilon.$
- (4)  $S \rightarrow (L) \mid a$  and  $L \rightarrow L, S \mid S.$

五、设含有否定  $\neg$  和蕴含  $\rightarrow$  运算的命题公式文法  $G(F)$  定义如下: (2020 年考题)

$$F \rightarrow F \rightarrow F \mid \neg F \mid (F) \mid a$$

其中: ‘a’, ‘ $\rightarrow$ ’, ‘ $\neg$ ’, ‘(’ 和 ‘)’ 为终结符,  $F$  是文法开始符号.

- (1) 试写出语句 “ $\neg(a \rightarrow a)$ ” 的一个最左推导;
- (2) 试消除文法  $G(F)$  中的左递归和左公因子;
- (3) 试对消除左递归后的文法所有非终结符求 First 集和 Follow 集;
- (4) 试对消除左递归后的文法构造 LL(1) 分析表, 从而说明消除左递归后的文法不是 LL(1) 文法;
- (5) 试利用你的分析表写出语句 “ $\neg(a \rightarrow a)$ ” 的一个正确的分析过程.

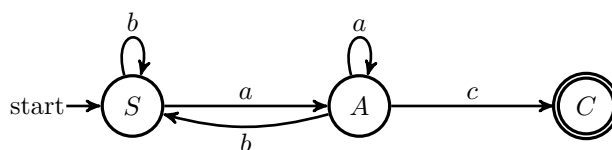
六、设文法  $G(F)$  如题二所示: (2020 年考题)

- (1) 试对语句 “ $\neg a \rightarrow a$ ” 画出两棵不同的语法树, 从而说明该文法为二义文法;
- (2) 试设计一个与文法  $G(F)$  等价的无二义的文法, 使得蕴含 ( $F \rightarrow F$ ) 为右结合, 且其的优先级低于否定.

七、设有文法  $G$  定义如下:

$$\begin{aligned} S &\rightarrow AB \\ A &\rightarrow aA \mid bA \mid \varepsilon \\ B &\rightarrow ac \end{aligned}$$

- (1) 试说明文法  $G$  不是 LL(1) 文法;  
解:  $\because \text{Select}\{A \rightarrow \varepsilon\} \supseteq \text{Follow}\{A\} = \{a\} \wedge \text{Select}\{A \rightarrow aB\} = \{a\}. \therefore \text{Select}\{A \rightarrow \varepsilon\} \cap \text{Select}\{A \rightarrow aB\} \neq \emptyset.$  这样文法  $G$  不是 LL(1) 文法.
- (2) 试设计一个与文法  $G$  等价的 LL(1) 文法.  
解:  $\because L(G) = L((a|b)^*ac)$ , 即  $L(G)$  是正则语言, 而识别正规表达式  $(a|b)^*ac$  的自动机如下所示:



则可验证其对应的正则文法是一个 LL(1) 文法:

$$\begin{aligned} S &\rightarrow aA \mid bS \\ A &\rightarrow aA \mid bS \mid c \end{aligned}$$

八、 设 C 语言的类型定义语句的文法如下所示：

$$\begin{aligned}
 D &\rightarrow T L \\
 T &\rightarrow \text{int} \mid \text{char} \\
 L &\rightarrow * L \quad (\text{指针}) \\
 &\mid L( ) \quad (\text{函数}) \\
 &\mid L[ ] \quad (\text{数组}) \\
 &\mid (L) \\
 &\mid \text{id}
 \end{aligned}$$

- (1) 写出语句 “`int * id[ ] ( )`” 两个不同的最右推导，和对应的两颗不同的语法树，从而说明该文法是二义文法；

解：

$  \begin{aligned}  &D \\  \xRightarrow{lm} &TL \\  \xRightarrow{lm} &\text{int } L \\  \xRightarrow{lm} &\text{int } *L \\  \xRightarrow{lm} &\text{int } *L( ) \\  \xRightarrow{lm} &\text{int } *L[ ] ( ) \\  \xRightarrow{lm} &\text{int } * \text{id} [ ] ( )  \end{aligned}  $	$\mid$	$  \begin{aligned}  &D \\  \xRightarrow{lm} &TL \\  \xRightarrow{lm} &\text{int } L \\  \xRightarrow{lm} &L( ) \\  \xRightarrow{lm} &L[ ] ( ) \\  \xRightarrow{lm} &\text{int } *L( ) \\  \xRightarrow{lm} &\text{int } * \text{id} [ ] ( )  \end{aligned}  $
--	--------	---

- (2) 改写该文法，使之消除二义性，并且运算的优先级别和 C 语言的类型定义一致，即括号运算优先级别最高，后缀运算优先级高于前缀运算.

解：引入表示后缀运算的非终结符  $A$  和运算量最小单位  $F$ ，则对应的等价无二义文法为：

$$\begin{aligned}
 D &\rightarrow T L \\
 T &\rightarrow \text{int} \mid \text{char} \\
 L &\rightarrow * L \mid A \\
 A &\rightarrow A( ) \mid A[ ] \mid F \\
 F &\rightarrow (L) \mid \text{id}
 \end{aligned}$$