

2019 级弘毅班《编译原理》第二次练习答案

一、 设有下述四组文法和其对应的终结符号串：

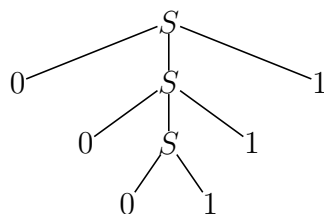
(a) (1)

$$\begin{aligned}
 & S \\
 \xRightarrow{lm} & 0S1 \\
 \xRightarrow{lm} & 00S11 \\
 \xRightarrow{lm} & 000111
 \end{aligned}$$

(2)

$$\begin{aligned}
 & S \\
 \xRightarrow{rm} & 0S1 \\
 \xRightarrow{rm} & 00S11 \\
 \xRightarrow{rm} & 000111
 \end{aligned}$$

(3)



(4) 该文法不是二义文法；

(5) $\{0^n 1^n \mid n \geq 1\}$.

(6) 无左递归和左公因子.

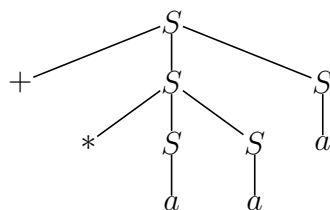
(b) (1)

$$\begin{aligned}
 & S \\
 \xRightarrow{lm} & +SS \\
 \xRightarrow{lm} & +*SSS \\
 \xRightarrow{lm} & +*aSS \\
 \xRightarrow{lm} & +*aaS \\
 \xRightarrow{lm} & +*aaa
 \end{aligned}$$

(2)

$$\begin{aligned} & S \\ \xRightarrow{rm} & + SS \\ \xRightarrow{rm} & + Sa \\ \xRightarrow{rm} & + *SSa \\ \xRightarrow{rm} & + *Saa \\ \xRightarrow{rm} & + *aaa \end{aligned}$$

(3)



- (4) 该文法不是二义文法;
- (5) 以字母 a 组成的前缀表达式.
- (6) 无左递归和左公因子.

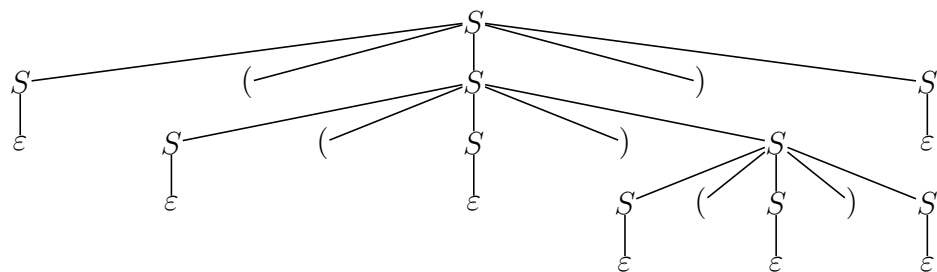
(c) (1) 两个不同的最左推导:

S		S
$\xRightarrow{lm} S(S)S$		$\xRightarrow{lm} S(S)S$
$\xRightarrow{lm} (S)S$		$\xRightarrow{lm} (S)S$
$\xRightarrow{lm} (\underline{S}(S)S)S \quad (S \rightarrow \varepsilon)$		$\xRightarrow{lm} (\underline{S}(S)S)S \quad (S \rightarrow S(S)S)$
$\xRightarrow{lm} ((S)S)S$		$\xRightarrow{lm} (S(S)S(S)S)S$
$\xRightarrow{lm} (()S)S$		$\xRightarrow{lm} ((S)S(S)S)S$
$\xRightarrow{lm} (()S(S)S)S$		$\xRightarrow{lm} (()S(S)S)S$
$\xRightarrow{lm} (() (S)S)S$		$\xRightarrow{lm} (() (S)S)S$
$\xRightarrow{lm} (() ()S)S$		$\xRightarrow{lm} (() ()S)S$
$\xRightarrow{lm} (() ())S$		$\xRightarrow{lm} (() ())S$
$\xRightarrow{lm} (()())$		$\xRightarrow{lm} (()())$

(2)

$$\begin{aligned}
 & S \\
 & \xRightarrow{rm} S(S)S \\
 & \xRightarrow{rm} S(S) \\
 & \xRightarrow{rm} S(S(S)S) \\
 & \xRightarrow{rm} S(S(S)S(S)S) \\
 & \xRightarrow{rm} S(S(S)S(S)) \\
 & \xRightarrow{rm} S(S(S)S()) \\
 & \xRightarrow{rm} S(S(S)()) \\
 & \xRightarrow{rm} S(S()()) \\
 & \xRightarrow{rm} S(()()) \\
 & \xRightarrow{rm} (()())
 \end{aligned}$$

(3)



- (4) 由 (A) 知该文法是二义文法；
 (5) 以 “(” 和 “)” 组成的嵌套的括号对。
 (6) 消除左递归后的文法：

$$\begin{aligned} S &\rightarrow S' \\ S' &\rightarrow (S)SS' \mid \varepsilon \end{aligned}$$

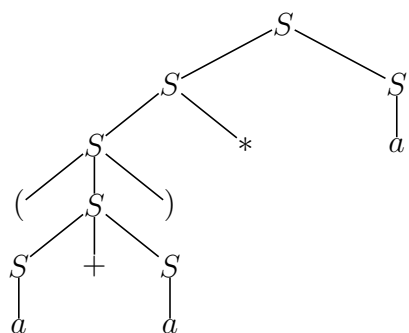
(d) (1)

$$\begin{aligned} &S \\ \Rightarrow &SS \\ \xRightarrow{lm} &S * S \\ \Rightarrow &(S) * S \\ \xRightarrow{lm} &(S + S) * S \\ \Rightarrow &(a + S) * S \\ \xRightarrow{lm} &(a + a) * S \\ \Rightarrow &(a + a) * a \end{aligned}$$

(2)

$$\begin{aligned} &S \\ \Rightarrow &SS \\ \xRightarrow{rm} &Sa \\ \Rightarrow &S * a \\ \xRightarrow{rm} &(S) * a \\ \Rightarrow &(S + S) * a \\ \xRightarrow{rm} &(S + a) * a \\ \Rightarrow &(a + a) * a \end{aligned}$$

(3)



(4) 该文法是二义文法. 因为 $a + a + a$ 有两个不同的最左推导:

$ \begin{aligned} &S \\ \xRightarrow{lm} &S + S \\ \xRightarrow{lm} &S + S + S \\ \xRightarrow{lm} &a + S + S \\ \xRightarrow{lm} &a + a + S \\ \xRightarrow{lm} &a + a + a \end{aligned} $		$ \begin{aligned} &S \\ \xRightarrow{lm} &S + S \\ \xRightarrow{lm} &a + S \\ \xRightarrow{lm} &a + S + S \\ \xRightarrow{lm} &a + a + S \\ \xRightarrow{lm} &a + a + a \end{aligned} $
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

(5) 以 a 组成的正规表达式, 其中并运算用 ‘+’ 表示.

(6) 消除左递归后的文法:

$$\begin{aligned}
 S &\rightarrow (S)S' \mid aS' \\
 S' &\rightarrow +SS' \mid SS' \mid *S' \mid \varepsilon
 \end{aligned}$$

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 ;

解: 该语言为正则语言, 其对应的正规表达式为: $(a|ab)^*$, 其对应的右线性文法为:

$$\begin{aligned}
 S &\rightarrow aS \mid aT \mid \varepsilon \\
 T &\rightarrow bS
 \end{aligned}$$

(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;

解: 该语言不是正则语言, 设 T 表示 a 和 b 出现相同的语言, 则任给一个 a 比 b 出现次数至少多一个的字符串 $w = (ta)^+t$, 其中 $t \in T$, 这样该语言的文法为:

$$\begin{aligned}
 S &\rightarrow TaS \mid TaT \\
 T &\rightarrow aTbT \mid bTaT \mid \varepsilon.
 \end{aligned}$$

文法为二义文法。

- (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/ $m > q$: ...)
 解: 设 $B = \{b^n c^p d^q \mid n = p+q\}$, $C = \{b^n c^p \mid n = p\}$ 和 $D = \{a^m b^n c^p \mid m+n = p\}$. 则

$$\begin{aligned} S &\rightarrow aSd \mid B \mid D \\ B &\rightarrow bBd \mid C \\ C &\rightarrow bCc \mid \varepsilon \\ D &\rightarrow aDc \mid C \end{aligned}$$

文法为二义文法, 因为 $\varepsilon \in B \cap D \subseteq S$. 该语言不是正则的.

- (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 $2p = 3q$.)
 解: 根据提示分析得到 a 增加 3 个与 b 增加 2 个保持平衡. 当 $n = 1$ 时, $m = 2$, 因此最短的语句为 aab , 故无二义文法为

$$S \rightarrow aaaSbb \mid aab$$

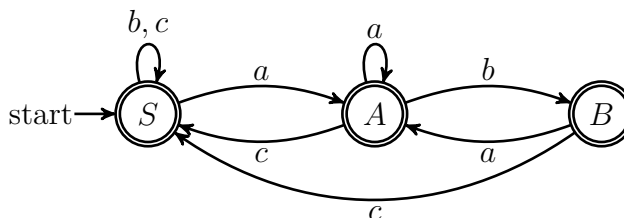
该语言不是正则的.

- (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.
 设 E 为 a 和 b 出现次数相等, A 为 a 比 b 多, B 为 b 比 a 多, S 为 a 和 b 不等, 则

$$\begin{aligned} S &\rightarrow A \mid B \\ A &\rightarrow AA \mid EaE \\ B &\rightarrow BB \mid EbE \\ E &\rightarrow aEbE \mid bEbE \mid \varepsilon \end{aligned}$$

二义文法, 非正则语言.

- (6) 所有的以 a 、 b 和 c 组成的字符串, 其中, 没有 abb 子串;
 解: 没有 abb 子串的 DFA 如下所示:



故对应的文法为

$$\begin{aligned} S &\rightarrow aA \mid bS \mid cS \mid \varepsilon \\ A &\rightarrow aA \mid bB \mid cS \mid \varepsilon \\ B &\rightarrow aA \mid cS \mid \varepsilon \end{aligned}$$

无二义文法, 该语言为正则语言.

(7) $\{a^m b^n \mid m > n \wedge m - n \text{ 是偶数}\}$ 。

解：设 $T = \{a^n b^n \mid n \in \mathbb{N}\}$ 。设 $w \in S$ ，则 $w = a^{2p} a^n b^n$ ($p \geq 1$)。这样可设计如下文法

$$\begin{aligned} S &\rightarrow aaS \mid aaT \\ T &\rightarrow aTb \mid \varepsilon \end{aligned}$$

无二义文法，该语言不是正则的。

三、 (1) 定义两非终结符集合：

$$\begin{aligned} \mathcal{E} &= \{U \mid U \in N \wedge U \xRightarrow{*} w \in T^*\} \\ \mathcal{A} &= \{U \mid U \in N \wedge S \xRightarrow{*} \alpha U \beta\} \end{aligned}$$

(1) 所有的经过 0 次或多次推出终结符的非终结符集 \mathcal{E} ：

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

\mathcal{E} 可递归定义为：

$$\begin{cases} \mathcal{E}_0 &= \emptyset \\ \mathcal{E}_{i+1} &= E_i \cup \{P \mid P \in N \wedge \exists P \rightarrow X_1 X_2 \cdots X_n \text{ 且 } X_j \in T \cup \mathcal{E}_i (j = 1, 2, \dots, n)\} \end{cases}$$

求 \mathcal{E} 的算法如下：

```
Set E = emptyset;
int change = 1;
while (change) {
    change = 0;
    foreach nonterminal P in (N - E) do
        foreach production (P -> X1 X2 ... Xn) do {
            int is_new = 1;
            foreach Xi do {
                if (Xi in (N - E)) {
                    is_new = 0;
                    break;
                }
            }
            if (is_new) {
                E = E + {P};
                change = 0;
            }
        }
    }
}
```

(2) 所有的从文法开始符号 S 能够推出的非终结符集 \mathcal{A} ：

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

\mathcal{A} 可递归定义为:

$$\begin{cases} \mathcal{A}_0 &= \{S\} \\ \mathcal{A}_{i+1} &= \mathcal{A}_i \cup \{X \mid X \in N \wedge \exists P \rightarrow \alpha X \beta \wedge P \in \mathcal{A}_i\} \end{cases} .$$

求 \mathcal{A} 的算法如下:

```
Set A = {S};
int change = 1;
while (change) {
    change = 0;
    foreach nonterminal P in (A) do
        foreach production (P -> X1 X2 ... Xn) do {
            foreach Xi do {
                if (Xi in (N - A)) {
                    change = 1;
                    A = A + {Xi};
                }
            }
        }
    }
}
```

算法:

- 1/ calcul the nonterminal set E which can generate the terminals.
- 2/ delete all production of nonterminal in $N - E$.
- 3/ calcul the accessible set A.
- 4/ delete all productions of nonterminal in $N - A$.

(2) 用下述文法检验你的算法:

$$\begin{aligned} S &\rightarrow 0 \mid A \\ A &\rightarrow AB \\ B &\rightarrow 1 \end{aligned}$$

- 1/ $\mathcal{E} = \{B, S\}$;
 - 2/ $N - \mathcal{E} = \{A\}$;
 - 3/ $\mathcal{A} = \{S\}$;
 - 4/ $(N - \mathcal{E}) - \mathcal{A} = \{B\}$.
- so $S \rightarrow 0$ 是仅剩的产生式.

思考: 为什么不能先计算 \mathcal{A} , 再计算 \mathcal{E} ?

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) $S \rightarrow SS+ \mid SS* \mid a$.

消除左递归和左公因子的等价文法:

$$\begin{aligned} S &\rightarrow aS' \\ S' &\rightarrow ST \mid \varepsilon \\ T &\rightarrow +S' \mid *S' \end{aligned}$$

该文法是 LL(1) 文法.

(2) $S \rightarrow 0S1 \mid 01$.

消除左递归和左公因子的等价文法:

$$\begin{aligned} S &\rightarrow 0S' \\ S' &\rightarrow S1 \mid 1 \end{aligned}$$

该文法是 LL(1) 文法. 其 LL(1) 分析表如下所示:

状态	0	1	\$
S	$S \rightarrow 0S'$		
S'	$S \rightarrow S1$	$S \rightarrow 1$	

(3) $S \rightarrow S(S)S \mid \varepsilon$.

消除左递归和左公因子的等价文法:

$$S \rightarrow (S)SS \mid \varepsilon$$

该文法是二义文法, 因此不是 LL(1) 文法. 不能用递归下降分析法分析.

状态	()	\$
S	$S \rightarrow (S)SS, S \rightarrow \varepsilon$	$S \rightarrow \varepsilon$	$S \rightarrow \varepsilon$

而与之等价的文法:

$$S \rightarrow (S)S \mid \varepsilon$$

是 LL(1) 文法. 其 LL(1) 分析表如下所示:

状态	()	\$
S	$S \rightarrow (S)SS$	$S \rightarrow \varepsilon$	$S \rightarrow \varepsilon$

(4) $S \rightarrow (L) \mid a$ and $L \rightarrow L, S \mid S$.

消除左递归和左公因子的等价文法:

$$\begin{aligned} S &\rightarrow (L) \mid a \\ L &\rightarrow SL' \\ L' &\rightarrow ,SL' \mid \varepsilon \end{aligned}$$

该文法是 LL(1) 文法.

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；
语句 “ $\neg(a \rightarrow a)$ ” 的最左推导如下：

$$\begin{array}{ll} F & \xRightarrow{lm} \neg F \\ & \xRightarrow{lm} \neg(F) \\ & \xRightarrow{lm} \neg(F \rightarrow F) \end{array} \quad \begin{array}{ll} & \xRightarrow{lm} \neg(a \rightarrow F) \\ & \xRightarrow{lm} \neg(a \rightarrow a) \end{array}$$

- (2) 试消除文法 $G(F)$ 中的左递归和左公因子；
消除左递归和左公因子后的文法如下：

$$\begin{array}{l} F \rightarrow \neg F F' \mid a F' \mid (F) F' \\ F' \rightarrow \rightarrow F F' \mid \varepsilon \end{array}$$

- (3) 试对消除左递归后的文法所有非终结符求 First 集和 Follow 集；

非终结符	First	Follow
F	a, (, \neg	\rightarrow ,), \$
F'	\rightarrow , ε	\rightarrow ,), \$

- (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；
LL(1) 分析表如下所示

	a	\neg	\rightarrow	()	\$
F	$F \rightarrow aF'$	$F \rightarrow \neg F F'$		$F \rightarrow (F)F'$		
F'			$F' \rightarrow \rightarrow F F' \mid \varepsilon$		$F' \rightarrow \varepsilon$	$F' \rightarrow \varepsilon$

分析表有冲突项，故不是 LL(1) 文法。

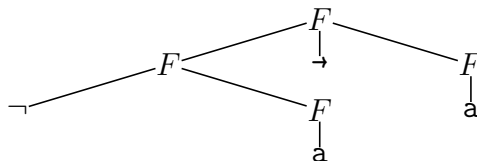
- (5) 试利用你的分析表写出语句 “ $\neg(a \rightarrow a)$ ” 的一个正确的分析过程。
语句 “ $\neg(a \rightarrow a)$ ” 的分析过程如下所示：

剩余串	分析栈	分析动作
$\neg(a \rightarrow a)\$$	$F\$$	$F \rightarrow \neg F F'$
$\neg(a \rightarrow a)\$$	$\neg F F' \$$	match-advance
$(a \rightarrow a)\$$	$F F' \$$	$F \rightarrow (F) F'$
$(a \rightarrow a)\$$	$(F) F' F' \$$	match-advance
$a \rightarrow a)\$$	$F) F' F' \$$	$F \rightarrow a F'$
$a \rightarrow a)\$$	$a F') F' F' \$$	match-advance
$\rightarrow a)\$$	$F') F' F' \$$	$F' \rightarrow \rightarrow F F'$
$\rightarrow a)\$$	$\rightarrow F F') F' F' \$$	match-advance
$a)\$$	$F F') F' F' \$$	$F \rightarrow a F'$
$a)\$$	$a F' F') F' F' \$$	match-advance
$)\$$	$F' F') F' F' \$$	$F' \rightarrow \varepsilon$
$)\$$	$F') F' F' \$$	$F' \rightarrow \varepsilon$
$)\$$	$) F' F' \$$	match-advance
$\$$	$F' F' \$$	$F' \rightarrow \varepsilon$
$\$$	$F' \$$	$F' \rightarrow \varepsilon$
$\$$	$\$$	分析成功

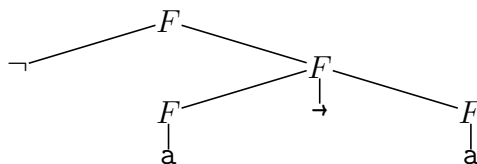
六、 设文法 $G(F)$ 如题二所示： (2020 年考题)

- (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；
语句 “ $\neg a \rightarrow a$ ” 的两棵不同的语法树为：

语法树 1:



语法树 2:



- (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。
无二义文法：

$$\begin{aligned}
 F &\rightarrow N \rightarrow F \mid N \\
 N &\rightarrow \neg N \mid A \\
 A &\rightarrow (F) \mid a
 \end{aligned}$$