

第 3 章 文法和语言

第 1 题

文法 $G = (\{A, B, S\}, \{a, b, c\}, P, S)$ 其中 P 为:

$S \rightarrow Ac|aB$

$A \rightarrow ab$

$B \rightarrow bc$

写出 $L(G[S])$ 的全部元素。

答案:

$L(G[S]) = \{abc\}$

第 2 题

文法 $G[N]$ 为:

$N \rightarrow D|ND$

$D \rightarrow 0|1|2|3|4|5|6|7|8|9$

$G[N]$ 的语言是什么?

答案:

$G[N]$ 的语言是 V^+ 。 $V = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

$N \Rightarrow ND \Rightarrow NDD \dots \Rightarrow NDDDD \dots D \Rightarrow D \dots D$

或者: 允许 0 开头的非负整数?

第 3 题

为只包含数字、加号和减号的表达式, 例如 $9-2+5$, $3-1$, 7 等构造一个文法。

答案:

$G[S]$:

$S \rightarrow S+D|S-D|D$

$D \rightarrow 0|1|2|3|4|5|6|7|8|9$

第 4 题

已知文法 $G[Z]$:

$Z \rightarrow aZb|ab$

写出 $L(G[Z])$ 的全部元素。

答案:

$$Z \Rightarrow aZb \Rightarrow aaZbb \Rightarrow aaa..Z...bbb \Rightarrow aaa..ab...bbb$$

$$L(G[Z]) = \{a^n b^n | n \geq 1\}$$

第 5 题

写一文法，使其语言是偶正整数的集合。 要求：

- (1) 允许 0 打头；
- (2) 不允许 0 打头。

答案:

(1) 允许 0 开头的偶正整数集合的文法

$$\begin{aligned} E &\rightarrow NT|D \\ T &\rightarrow NT|D \\ N &\rightarrow D|1|3|5|7|9 \\ D &\rightarrow 0|2|4|6|8 \end{aligned}$$

(2) 不允许 0 开头的偶正整数集合的文法

$$\begin{aligned} E &\rightarrow NT|D \\ T &\rightarrow FT|G \\ N &\rightarrow D|1|3|5|7|9 \\ D &\rightarrow 2|4|6|8 \\ F &\rightarrow N|0 \\ G &\rightarrow D|0 \end{aligned}$$

第 6 题

已知文法 G:

$$\langle \text{表达式} \rangle ::= \langle \text{项} \rangle \mid \langle \text{表达式} \rangle + \langle \text{项} \rangle$$

$$\langle \text{项} \rangle ::= \langle \text{因子} \rangle \mid \langle \text{项} \rangle * \langle \text{因子} \rangle$$

$$\langle \text{因子} \rangle ::= (\langle \text{表达式} \rangle) \mid i$$

试给出下述表达式的推导及语法树。

(5) $i + (i + i)$

(6) $i + i * i$

答案:

(5) <表达式>

=><表达式>+<项>

=><表达式>+<因子>

=><表达式>+ (<表达式>)

=><表达式>+ (<表达式>+<项>)

=><表达式>+ (<表达式>+<因子>)

=><表达式>+ (<表达式>+i)

=><表达式>+ (<项>+i)

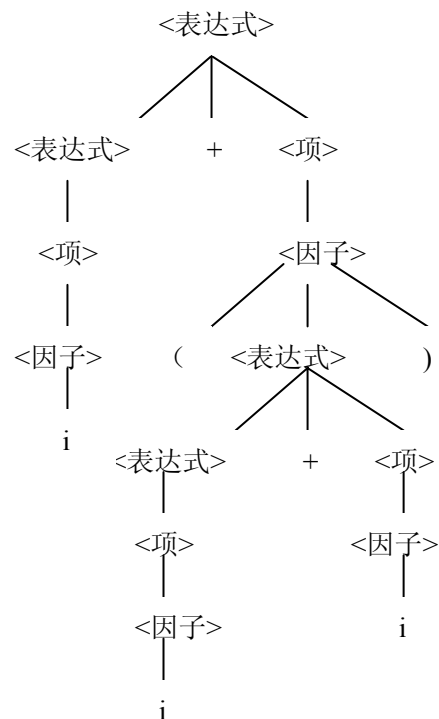
=><表达式>+ (<因子>+i)

=><表达式>+ (i+i)

=><项>+ (i+i)

=><因子>+ (i+i)

=>i+ (i+i)



(6) <表达式>

=><表达式>+<项>

=><表达式>+<项>*<因子>

=><表达式>+<项>*i

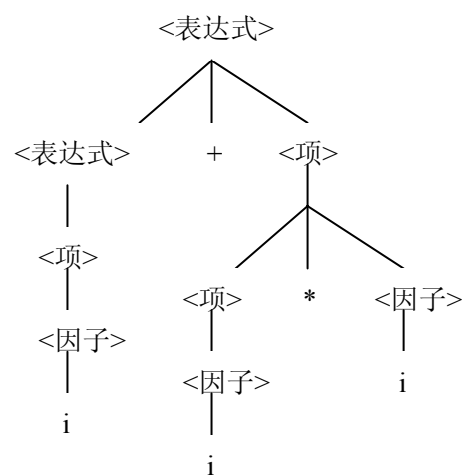
=><表达式>+<因子>*i

=><表达式>+i*i

=><项>+i*i

=><因子>+i*i

=>i+i*i



第 7 题

证明下述文法 $G[\langle \text{表达式} \rangle]$ 是二义的。

$\langle \text{表达式} \rangle ::= a[(\langle \text{表达式} \rangle)] \langle \text{表达式} \rangle \langle \text{运算符} \rangle \langle \text{表达式} \rangle$

$\langle \text{运算符} \rangle ::= +|-|*|/$

答案:

可为句子 $a+a*a$ 构造两个不同的最右推导:

最右推导 1 $\langle \text{表达式} \rangle \Rightarrow \langle \text{表达式} \rangle \langle \text{运算符} \rangle \langle \text{表达式} \rangle$

$\Rightarrow \langle \text{表达式} \rangle \langle \text{运算符} \rangle a$

$\Rightarrow \langle \text{表达式} \rangle * a$

$\Rightarrow \langle \text{表达式} \rangle \langle \text{运算符} \rangle \langle \text{表达式} \rangle * a$

$\Rightarrow \langle \text{表达式} \rangle \langle \text{运算符} \rangle a * a$

$\Rightarrow \langle \text{表达式} \rangle + a * a$

$\Rightarrow a + a * a$

最右推导 2 $\langle \text{表达式} \rangle \Rightarrow \langle \text{表达式} \rangle \langle \text{运算符} \rangle \langle \text{表达式} \rangle$

$\Rightarrow \langle \text{表达式} \rangle \langle \text{运算符} \rangle \langle \text{表达式} \rangle \langle \text{运算符} \rangle \langle \text{表达式} \rangle$

$\Rightarrow \langle \text{表达式} \rangle \langle \text{运算符} \rangle \langle \text{表达式} \rangle \langle \text{运算符} \rangle a$

$\Rightarrow \langle \text{表达式} \rangle \langle \text{运算符} \rangle \langle \text{表达式} \rangle * a$

$\Rightarrow \langle \text{表达式} \rangle \langle \text{运算符} \rangle a * a$

$\Rightarrow \langle \text{表达式} \rangle + a * a$

$\Rightarrow a + a * a$

第 8 题

文法 $G[S]$ 为:

$S \rightarrow Ac|aB$

$A \rightarrow ab$

$B \rightarrow bc$

该文法是否为二义的? 为什么?

答案:

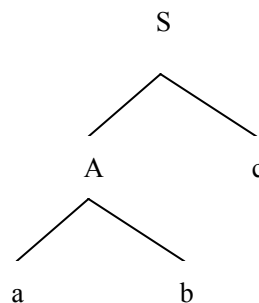
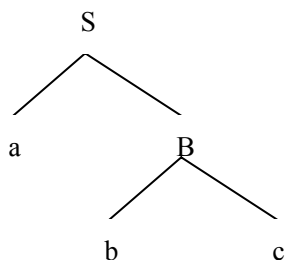
对于串 abc

(1) $S \Rightarrow Ac \Rightarrow abc$ (2) $S \Rightarrow aB \Rightarrow abc$

即存在两不同的最右推导。所以, 该文法是二义的。

或者:

对输入字符串 abc , 能构造两棵不同的语法树, 所以它是二义的。



第 9 题

考虑下面上下文无关文法:

$S \rightarrow SS^*|SS+|a$

(1) 表明通过此文法如何生成串 $aa+a^*$, 并为该串构造语法树。

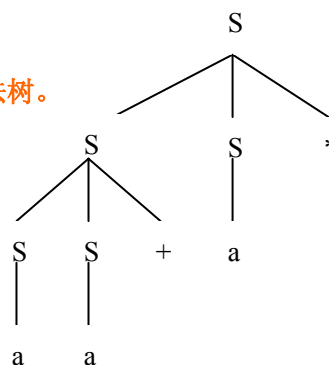
(2) $G[S]$ 的语言是什么?

答案:

(1) 此文法生成串 $aa+a^*$ 的最右推导如下

$S \Rightarrow SS^* \Rightarrow SS^* \Rightarrow Sa^* \Rightarrow SS+a^* \Rightarrow Sa+a^* \Rightarrow aa+a^*$

(2) 该文法生成的语言是: $*$ 和 $+$ 的后缀表达式, 即逆波兰式。



第 10 题

文法 $S \rightarrow S(S)S | \epsilon$

(1) 生成的语言是什么？

(2) 该文法是二义的吗？说明理由。

答案：

(1) 嵌套的括号

(2) 是二义的，因为对于 $()()$ 可以构造两棵不同的语法树。

第 11 题

令文法 $G[E]$ 为：

$E \rightarrow T | E+T | E-T$

$T \rightarrow F | T * F | T / F$

$F \rightarrow (E) | i$

证明 $E+T * F$ 是它的一个句型，指出这个句型的所有短语、直接短语和句柄。

答案：

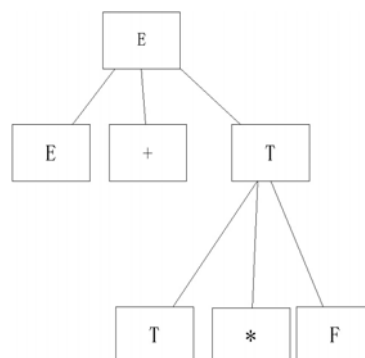
此句型对应语法树如右，故为此文法一个句型。

或者：因为存在推导序列： $E \Rightarrow E+T \Rightarrow E+T * F$ ，所以 $E+T * F$ 句型

此句型相对于 E 的短语有： $E+T * F$ ；相对于 T 的短语有 $T * F$

直接短语为： $T * F$

句柄为： $T * F$



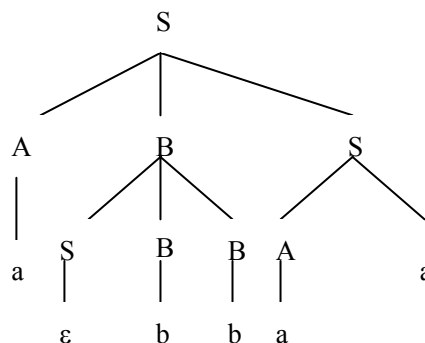
第 13 题

一个上下文无关文法生成句子 $abb a a$ 的推导树如下：

(1) 给出串 $abb a a$ 最左推导、最右推导。

(2) 该文法的产生式集合 P 可能有哪些元素？

(3) 找出该句子的所有短语、直接短语、句柄。



答案:

(1)串 abbaa 最左推导:

$S \Rightarrow ABS \Rightarrow aBS \Rightarrow aSBB \Rightarrow aBBS \Rightarrow abBS \Rightarrow abbS \Rightarrow abbAa \Rightarrow abbaa$

最右推导:

$S \Rightarrow ABS \Rightarrow ABaA \Rightarrow ABaa \Rightarrow ASBBaa \Rightarrow ASBbaa \Rightarrow ASbbaa \Rightarrow Abbaa \Rightarrow abbaa$

(2)产生式有: $S \rightarrow ABS \mid Aa \mid \varepsilon$ $A \rightarrow a$ $B \rightarrow SBB \mid b$

可能元素有: ε aa ab abbaa aaabbaa

(3)该句子的短语有:

a 是相对 A 的短语

ε 是相对 S 的短语

b 是相对 B 的短语

ebb 是相对 B 的短语

aa 是相对 S 的短语

aεbbaa 是相对 S 的短语

直接短语有: a ε b

句柄是: a

第 14 题

给出生成下述语言的上下文无关文法:

(1) $\{a^n b^n a^m b^m \mid n, m \geq 0\}$

(2) $\{1^n 0^m 1^m 0^n \mid n, m \geq 0\}$

(3) $\{WaWr \mid W \text{ 属于 } \{0|a\}^*, Wr \text{ 表示 } W \text{ 的逆}\}$

答案:

(1)

$S \rightarrow AA$

$A \rightarrow aAb \mid \varepsilon$

(2)

$S \rightarrow 1S0 \mid A$

$A \rightarrow 0A1 \mid \varepsilon$

(3)

$S \rightarrow 0S0 \mid 1S1 \mid \varepsilon$

第 16 题

给出生成下述语言的三型文法：

(1) $\{a^n | n \geq 0\}$

(2) $\{a^n b^m | n, m \geq 1\}$

(3) $\{a^n b^m c^k | n, m, k \geq 0\}$

答案：

(1) $S \rightarrow aS | \varepsilon$

(2)

$S \rightarrow aA$

$A \rightarrow aA | B$

$B \rightarrow bB | b$

(3)

$A \rightarrow aA | B$

$B \rightarrow bB | C$

$C \rightarrow cC | \varepsilon$

第 17 题

习题 7 和习题 11 中的文法等价吗？

答案：

等价。

第 18 题

解释下列术语和概念：

(1) 字母表

(2) 串、字和句子

(3) 语言、语法和语义

答案：

(1) 字母表：是一个非空有穷集合。

(2) 串：符号的有穷序列。

字：字母表中的元素。

句子：如果 $Z \xRightarrow{+} x, x \in V^*T$ 则称 x 是文法 G 的一个句子。

(3) 语言：它是由句子组成的集合，是由一组记号所构成的集合。程序设计的语言就是所有该语言的程序的全体。语言可以看成在一个基本符号集上定义的，按一定规则构成的一切基本符号串组成的集合。

语法：表示构成语言句子的各个记号之间的组合规律。程序的结构或形式。

语义：表示按照各种表示方法所表示的各个记号的特定含义。语言所代表的含义。

附加题

问题 1:

给出下述文法所对应的正规式:

$S \rightarrow 0A|1B$

$A \rightarrow 1S|1$

$B \rightarrow 0S|0$

答案:

$$R = (01 | 10) (01 | 10)^*$$

问题 2:

已知文法 $G[A]$, 写出它定义的语言描述

$G[A]: A \rightarrow 0B|1C$

$B \rightarrow 1|1A|0BB$

$C \rightarrow 0|0A|1CC$

答案:

$G[A]$ 定义的语言由 0、1 符号串组成, 串中 0 和 1 的个数相同.

问题 3:

给出语言描述, 构造文法.

构造一文法,其定义的语言是由算符+, *, (,)和运算对象 a 构成的算术表达式的集合.

答案一:

$G[E] \ E \rightarrow E+T|T$

$T \rightarrow T^* F|F$

$F \rightarrow (E)|a$

答案二:

$G[E] \ E \rightarrow E+E|E^* E|(E)|a$

问题 4:

已知文法 $G[S]$:

$S \rightarrow dAB$

$$A \rightarrow aA|a$$

$$B \rightarrow \varepsilon | bB$$

相应的正规式是什么？G[S]能否改写成为等价的正规文法？

答案：

正规式是 daa^*b^* ；

相应的正规文法为(由自动化简来)：

G[S]: $S \rightarrow dA$ $A \rightarrow a|aB$ $B \rightarrow aB|a|b|bC$ $C \rightarrow bC|b$

也可为(观察得来): G[S]: $S \rightarrow dA$ $A \rightarrow a|aA|aB$ $B \rightarrow bB|\varepsilon$

问题 5:

已知文法 G:

$$E \rightarrow E+T|E-T|T$$

$$T \rightarrow T*F|T/F|F$$

$$F \rightarrow (E)|i$$

试给出下述表达式的推导及语法树

(1) i ;

(2) $i*i+i$

(3) $i+i*i$

(4) $i+(i+i)$

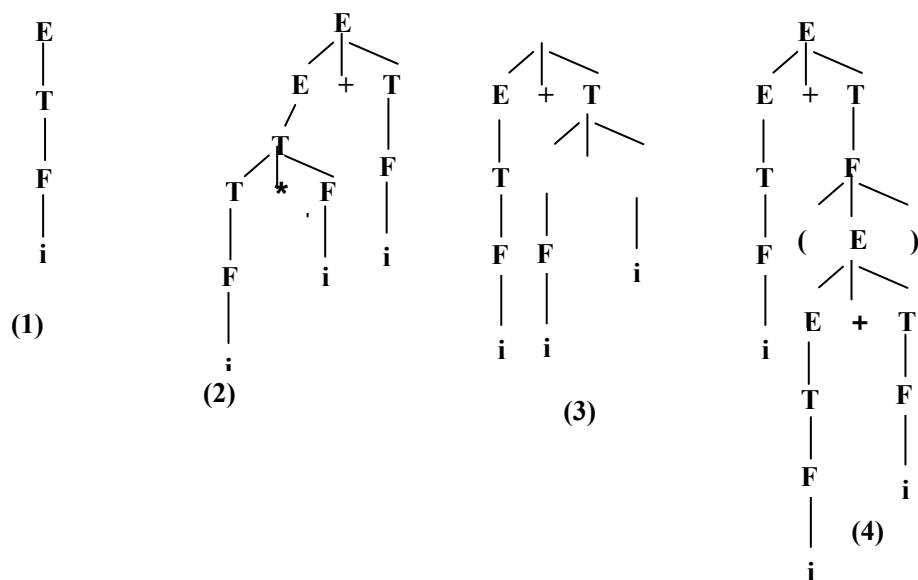
答案:

(1) $E \Rightarrow T \Rightarrow F \Rightarrow i$

(2) $E \Rightarrow E+T \Rightarrow T+T \Rightarrow T*F+T \Rightarrow F*F+T \Rightarrow i*F+T \Rightarrow i*i+T \Rightarrow i*i+F \Rightarrow i*i+i$

(3) $E \Rightarrow E+T \Rightarrow T+T \Rightarrow F+T \Rightarrow i+T \Rightarrow i+T*F \Rightarrow i+F*F \Rightarrow i+i*F \Rightarrow i+i*i$

(4) $E \Rightarrow E+T \Rightarrow T+T \Rightarrow F+T \Rightarrow i+T \Rightarrow i+F \Rightarrow i+(E) \Rightarrow i+(E+T) \Rightarrow i+(T+T) \Rightarrow i+(F+T) \Rightarrow i+(i+T) \Rightarrow i+(i+F) \Rightarrow i+(i+i)$



问题 6:

已知文法 $G[E]$:

$$E \rightarrow ET + |T$$

T → TF* | F

$$\mathbf{F} \rightarrow \mathbf{F}^{\wedge} \mid \mathbf{a}$$

试证: FF^* 是文法的句型, 指出该句型的短语、简单短语和句柄.

答案:

该句型对应的语法树如下:

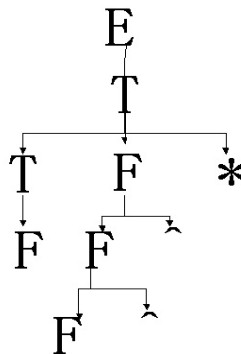
该句型相对于 E 的短语有 FF^{**}

相对于 T 的短语有 FF^*, F

相对于 F 的短语有 $F^{\wedge}; F^{\wedge\wedge}$

简单短语有 $F; F^{\wedge}$

句柄为 F.



问题 7:

适当变换文法，找到下列文法所定义语言的一个无二义的文法：

$$S \rightarrow SaS \mid SbS \mid ScS \mid d$$

答案:

该文法的形式很典型,可以先采用优先级联规则变换文法,然后再规定结合性对文法做进一步变换,即可消除二义性。

设 a 、 b 和 c 的优先级别依次增高, 根据优先级联规则将文法变换为:

$$S \rightarrow SaS \mid A$$
$$A \rightarrow AbA \mid C$$
$$C \rightarrow CcC \mid d$$

规定结合性为左结合，进一步将文法变换为：

$$S \rightarrow SaA \mid A$$
$$A \rightarrow AbC \mid C$$
$$C \rightarrow CcF \mid F$$
$$F \rightarrow d$$

该文法为非二义的。

问题 8:

构造产生如下语言的上下文无关文法:

$$(1) \{a^n b^{2n} c^m \mid n, m \geq 0\}$$

$$(2) \{a^n b^m c^{2m} \mid n, m \geq 0\}$$

$$(3) \{a^m b^n \mid m \geq n\}$$

$$(4) \{a^m b^n c^p d^q \mid m+n=p+q\}$$

$$(5) \{uawb \mid u, w \in \{a, b\}^* \wedge |u| = |w|\}$$

答案:

(1) 根据上下文无关文法的特点, 要产生形如 $a^n b^{2n} c^m$ 的串, 可以分别产生形如 $a^n b^{2n}$ 和形如 c^m 的串。设计好的文法是否就是该语言的文法? 严格地说, 应该给出证明。但若不是特别指明, 通常可以忽略这一点。

对于该语言, 存在一个由以下产生式定义的上下文无关文法 $G[S]$:

$$\begin{aligned} S &\rightarrow AB \\ A &\rightarrow \varepsilon \mid aAbb \\ B &\rightarrow \varepsilon \mid cB \end{aligned}$$

(2) 同样, 要产生形如 $a^n b^m c^{2m}$ 的串, 可以分别产生形如 a^n 和形如 $b^m c^{2m}$ 的串。对于该语言, 存在一个由以下产生式定义的上下文无关文法 $G[S]$:

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

(3) 考虑在先产生同样数目的 a, b , 然后再生成多余的 a 。以下 $G[S]$ 是一种解法:

$$S \rightarrow aSb \mid aS \mid \varepsilon$$

(4) 以下 $G[S]$ 是一种解法:

$$\begin{aligned} S &\rightarrow aSd \mid A \mid D \\ A &\rightarrow bAd \mid B \\ D &\rightarrow aDc \mid B \\ B &\rightarrow bBc \mid \varepsilon \end{aligned}$$

注: a 不多于 d 时, b 不少于 c ; 反之, a 不少于 d 时, b 不多于 c 。前一种情形通过对应 A , 后一种情形对应 D 。

(5) 以下 $G[S]$ 是一种解法:

$$\begin{aligned} S &\rightarrow Ab \\ A &\rightarrow BAB \mid a \end{aligned}$$

$$B \rightarrow a \mid b$$

问题 9:

下面的文法 $G(S)$ 描述由命题变量 p 、 q ，联结词 \wedge （合取）、 \vee （析取）、 \neg （否定）构成的命题公式集合：

$$\begin{aligned} S &\rightarrow S \vee T \mid T \\ T &\rightarrow T \wedge F \mid F \\ F &\rightarrow \neg F \mid p \mid q \end{aligned}$$

试指出句型 $\neg F \vee \neg q \wedge p$ 的直接短语（全部）以及句柄。

答案:

直接短语: $p, q, \neg F$

句柄: $\neg F$

问题 10:

设字母表 $A=\{a\}$ ，符号串 $x=aaa$ ，写出下列符号串及其长度: x^0, xx, x^5 以及 A^+ 。

答案:

$$x^0=(aaa)^0=\varepsilon \quad |x^0|=0$$

$$xx=aaaaaa \quad |xx|=6$$

$$x^5=aaaaaaaaaaaaaaa \quad |x^5|=15$$

$$A^+=A^1 \cup A^2 \cup \dots \cup A^n \cup \dots = \{a, aa, aaa, aaaa, aaaaa, \dots\}$$

$$A^*=A^0 \cup A^1 \cup A^2 \cup \dots \cup A^n \cup \dots = \{\varepsilon, a, aa, aaa, aaaa, aaaaa, \dots\}$$

问题 11:

令 $\Sigma=\{a, b, c\}$ ，又令 $x=abc$ ， $y=b$ ， $z=aab$ ，写出如下符号串及它们的长度: xy ， xyz ， $(xy)^3$

答案:

$$xy=abcb \quad |xy|=4$$

$$xyz=abcbaab \quad |xyz|=7$$

$$(xy)^3=(abcb)^3=abcbabcbabcb \quad |(xy)^3|=12$$

问题 12:

已知文法 $G[Z]$: $Z ::= U0 \mid V1$ 、 $U ::= Z1 \mid 1$ 、 $V ::= Z0 \mid 0$ ，请写出全部由此文法描述的只含有四个符号的句子。

答案:

$Z \Rightarrow U0 \Rightarrow Z10 \Rightarrow U010 \Rightarrow 1010$

$Z \Rightarrow U0 \Rightarrow Z10 \Rightarrow V110 \Rightarrow 0110$

$Z \Rightarrow V1 \Rightarrow Z00 \Rightarrow U000 \Rightarrow 1000$

$Z \Rightarrow V1 \Rightarrow Z00 \Rightarrow V100 \Rightarrow 0100$

问题 13:

已知文法 $G[S]$: $S ::= AB$ $A ::= aA \mid \varepsilon$ $B ::= bBc \mid bc$, 写出该文法描述的语言。

答案:

$A ::= aA \mid \varepsilon$ 描述的语言: $\{a^n | n \geq 0\}$

$B ::= bBc \mid bc$ 描述的语言: $\{b^n c^n | n \geq 1\}$

$L(G[S]) = \{a^n b^m c^m | n \geq 0, m \geq 1\}$

问题 14:

已知文法 $E ::= T \mid E+T \mid E-T$ 、 $T ::= F \mid T * F \mid T / F$ 、 $F ::= (E) \mid i$, 写出该文法的开始符号、终结符号集合 V_T 、非终结符号集合 V_N 。

答案:

开始符号: E

$V_T = \{+, -, *, /, (,), i\}$

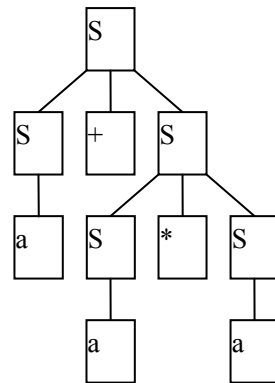
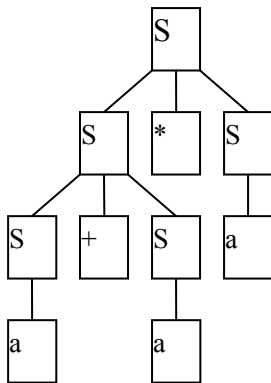
$V_N = \{E, F, T\}$

问题 15:

设有文法 $G[S]$: $S ::= S * S \mid S + S \mid (S)$, 该文法是二义性文法吗?

答案:

根据所给文法推导出句子 $a+a*a$, 画出了两棵不同的语法树, 所以该文法是二义性文法。



问题 16:

写一文法，使其语言是奇正整数集合。

答案:

$A ::= 1|3|5|7|9|NA$

$N ::= N0|N1|N2|N3|N4|N5|N6|N7|N8|N9|$

$N ::= 0|1|2|3|4|5|6|7|8|9$