形式语言与自动机期末速通

2. 文法

2.1 文法的形式定义

[**定义2.1.1**] 文法G由四元组(V,T,P,S)定义,记作$G=(V,T,P,S)$,其中:

(1)V为变量的非空有穷集.$\forall A\in V$称为一个**语法变量**,简称**变量**,也称**非终极符**.

V中的变量表示语法范畴,故有时也称V为**语法范畴**.

(2)T为终极符的集合.$\forall a\in T$称为**终极符**.

T中的字符是语言句子中出现的字符,则$V\bigcap T=\varnothing$.

(3)P为产生式的非空有穷集.P中的元素有形式$\alpha\rightarrow\beta$,称为**产生式**或**定义式**或**语法规则**,其中$\alpha\in(V\bigcup T)^+$且α中至少有V中元素的一个出现,$\beta\in(V\bigcup T)^*$.α称为产生式的**左部**,β称为产生式的**右部**.

(4)S为文法G的**开始符号**,则$S\subseteq V$.

[**注1**] 对一组有相同左部的产生式$\alpha\rightarrow\beta_1,\cdots,\alpha\rightarrow\beta_n$,简记为$\alpha\rightarrow\beta_1|\cdots|\beta_n$,其中$\beta_1,\cdots,\beta_n$称为**候选式**.

[**注2**]

①用字母表中较靠前的大写字母表示语法变量,如$A,B,C,\cdots$.

②用字母表中较靠后的大写字母表示语法变量或终极符,如$X,Y,Z,\cdots$.

③用字母表中较靠前的小写字母表示终极符,如$a,b,c,\cdots$.

④用字母表中较靠后的小写字母表示终极符构成的句子,如$x,y,z,\cdots$.

⑤用希腊字母表示由语法变量和终极符构成的句子,如$\alpha,\beta,\gamma,\cdots$.

[**注3**] 对一个文法,可只列出其所有产生式,且第一个产生式的左部是该文法的开始符号.

[**例2.1.1**] 下列四元组都是文法:

(1)$(\{A\},\{0,1\},\{A\rightarrow 01,A\rightarrow 0A1,A\rightarrow 1A0\},A)$.

(2)$(\{A,B\},\{0,1\},\{A\rightarrow 0,A\rightarrow 1,A\rightarrow 0A,A\rightarrow 1A\},A)$.

(3)$\begin{matrix}(\{S,A,B,C,D\},\{a,b,c,d,\#\},\\ \{S\rightarrow ABCD,S\rightarrow abc\#,A\rightarrow aaA,BC\rightarrow bccC,cC=cccC,CD\rightarrow d\#\},S)\end{matrix}$.

(4)$\begin{matrix}(\{A,B,C,D,E,F,S\},\{a,b,c,e\},\\ \{S\rightarrow ABC|abc,D\rightarrow e|a,FB\rightarrow c,A\rightarrow A,E\rightarrow abc|\varepsilon\})\end{matrix}$.

[**定义2.1.2**] 对文法$G=(V,T,P,S)$,若产生式$(\alpha\rightarrow\beta)\in P$,则对行$\gamma,\delta\in(V\bigcup T)^*$,称句子$\gamma\alpha\delta$在$G$中**直接推导**出句子$\gamma\beta\delta$,记作$\gamma\alpha\delta\Rightarrow_G\gamma\beta\delta$.直接推导简称**推导**,又称**派生**.反过来,称句子$\gamma\beta\delta$在G中**直接归约**为句子$\gamma\alpha\delta$.不强调归约的直接性时,直接归约可简称**归约**.

[**定义2.1.3**] 对文法$G=(V,T,P,S)$,归约$\Rightarrow_G$、$\Rightarrow_G^+$、$\Rightarrow_G^*$是集合$(V\bigcup T)^*$上的二元关系,其中:

(1)$\alpha \Rightarrow_G^n \beta$表示句子$\alpha$在$G$中经$n$步推导出句子$\beta$,$\beta$在$G$中经$n$步归约为$\alpha$,

即∃行$\alpha_1,\cdots,\alpha_{n-1}\in(V\bigcup T)^*\ s.t.\ \alpha\Rightarrow_G\alpha_1,\cdots,\alpha_1\Rightarrow_G\alpha_2,\cdots,\alpha_{n-1}\Rightarrow_G\beta$.

$n=0$时,有$\alpha=\beta$,即$\alpha\Rightarrow_G^0\alpha$.

(2)$\alpha \Rightarrow_G^+ \beta$表示句子$\alpha$在$G$中经至少1步推导出句子$\beta$,$\beta$在$G$中经至少1步归约为$\alpha$.

(3)$\alpha \Rightarrow_G^* \beta$表示句子$\alpha$在$G$中步推导若干步推导出句子$\beta$,$\beta$在$G$中经若干步归约为$\alpha$.

[**注**] 实际应用中常省略G,即用$\Rightarrow,\Rightarrow^n,\Rightarrow^+,\Rightarrow^*$分别代替$\Rightarrow_G,\Rightarrow_G^n,\Rightarrow_G^+,\Rightarrow_G^*$.

[**例2.1.2**] 设文法$G=(\{A\},\{a\},\{A\to a|aA\},A)$,则句子$aaaaa$的推导过程:

$A\Rightarrow aA$ *使用产生式$A\to aA$.

$\Rightarrow aaA$ *使用产生式$A\to aA$.

$\Rightarrow aaaA$ *使用产生式$A\to aA$.

$\Rightarrow aaaaA$ *使用产生式$A\to aA$.

$\Rightarrow aaaaa$ *使用产生式$A\to a$.

[**例2.1.3**] 设文法$G=(\{S,A,B\},\{0,1\},\{S\rightarrow A|AB. A\rightarrow 0|0A,B\rightarrow 1|11\},S)$.

(1)$A\Rightarrow^n 0^n$ *先连续使用$(n-1)$次产生式$A\rightarrow 0A$,再使用1次产生式$A\rightarrow 0$.

(2)$B\Rightarrow 1$ *使用产生式$B\rightarrow 1$.

(3)$B\Rightarrow 11$ *使用产生式$B\rightarrow 11$.

(4)语法范畴A表示的集合$L(A)=\{0,00,000,\cdots\}=\{0^n\mid n\geq 1\}$.

语法范畴B表示的集合$L(B)=\{1,11\}$.

语法范畴S表示的集合$L(S)=L(A)\bigcup L(A)L(B)=\{0^n\mid n\geq 1\}\bigcup\{0^n\mid n\geq 1\}\{1,11\}$.

[注] 设字母表Σ,句子$x,y\in\Sigma^+$.

(1)欲使语法范畴$D=\{x^n\mid n\geq 0\}$,可用产生式组$D\to\varepsilon|xD$.

(2)欲使语法范畴$D=\{x^ny^n\mid n\geq 1\}$,可用产生式组$D\to xy|xDy$.

(3)欲使语法范畴$D=\{x^ny^n\mid n\geq 0\}$,可用产生式组$D\to\varepsilon|xDy$.

[定义2.1.4] 对文法$G=(V,T,P,S)$,定义其**语言**$L(G)=\{w\mid w\in T^*\wedge S\Rightarrow^* w\}$,其中$\forall w\in L(G)$称为$G$产生的一个**句子**.对$\forall$句子$\alpha\in(V\bigcup T)^*$,若$S\Rightarrow^*=\alpha$,则称$\alpha$是$G$产生的一个**句型**.

[注]

(1)句子是从S开始的,在G中可推导出的终极符行,不含语法变量.

(2)句型是从S开始的,在G中可推搭配出来的符号行,可能含语法变量.

2.2 文法的构造

[例2.2.1] 构造产生C语言的合法标识符的文法.

[解] $$\begin{aligned}G=(&\{<标识符>,<大写字母>,<小写字母>,<阿拉伯数字>,<下划线>\},\\&\{0,\cdots,9,A,\cdots,Z,a,\cdots,z,\_\},P,<标识符>)\end{aligned},$$

其中$$\begin{aligned}P=\{&<标识符>\to<大写字母>\mid<小写字母>,<标识符>\to<标识符><大写字母>|\\&<标识符><小写字母>\mid<标识符><阿拉伯数字>\mid<标识符><下划线>\}\end{aligned},$$

$<大写字母>\to A|B|C|D|E|F|G|H|I|J|K|L|M|N|O|P|Q|R|S|T|U|V|W|X|Y|Z,$

$<小写字母>\to a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z,$

$<阿拉伯数字>\to 0|1|2|3|4|5|6|7|8|9,<下划线>\to\_.$

[注]

(1)不可简写为$<大写字母>\to A|\cdots|Z$.

(2)不可用$A\to a^3$表示$A\to aaa$.

(3)不可用$A\to a^n$表示A可产生任意多个a.

[例2.2.2] 文法G的产生式子如下:①$S\to aBC|aSBC$;②$aB\to ab$;③$bB\to bb$;④$CB\to BC$;⑤$bC\to bc$;⑥$cC\to cc$.给出句子$aaabbbccc$的两个不同的推导和归约.

[解]

(1)$S\Rightarrow aSBC$

$\Rightarrow aaSBCBC$

$\Rightarrow aaaBCBCBC$

$\Rightarrow aaabCBCBC$

$\Rightarrow aaabBCCBC$

$\Rightarrow aaabbCCBC$

$\Rightarrow aaabbCBCC$

$\Rightarrow aaabbBCCC$

$\Rightarrow aaabbbCCC$

$\Rightarrow aaabbbcCC$

$\Rightarrow aaabbbccC$

$\Rightarrow aaabbbccc$

(2)$S\Rightarrow aSBC$

$\Rightarrow aaSBCBC$

$\Rightarrow aaaBCBCBC$

$\Rightarrow aaaBBCCBC$

$\Rightarrow aaaBBCBCC$

$\Rightarrow aaaBBBCCC$

$\Rightarrow aaabBBCCC$

$\Rightarrow aaabbBCCC$

$\Rightarrow aaabbbCCC$

$\Rightarrow aaabbbcCC$

$\Rightarrow aaabbbccC$

$\Rightarrow aaabbbccc$

[**例2.2.3**] 构造文法G,使得其语言$L(G)=\{0,1,00,11\}$.

[**解**]

(1)先将文法的开始符号定义为$L(G)$中的四个句子:

$G_1=(\{S\},\{0,1\},\{S\rightarrow 0,S\rightarrow 1,S\rightarrow 00,S\rightarrow 11\},S)$.

(2)用变量A表示0,变量B表示1.

$G_2=(\{S,A,B\},\{0,1\},\{S\rightarrow A,S\rightarrow B,S\rightarrow AA,S\rightarrow BB,A\rightarrow 0,B\rightarrow 1\},S)$.

(3)将G_2化简:

$G_3=(\{S,A,B\},\{0,1\},\{S\rightarrow 0,S\rightarrow 1,S\rightarrow 0A,S\rightarrow 1B,A\rightarrow 0,B\rightarrow 1\},S)$.

(4)可在V、T中加入其他元素得到形式上不同的文法,但它们定义的语言都相同.

$$G_4=(\{S,A,B,C\},\{0,1,2\},\\ \{S\rightarrow A,S\rightarrow B,S\rightarrow AA,S\rightarrow BB,A\rightarrow 0,B\rightarrow 1,CACS\rightarrow 21,C\rightarrow 11,C\rightarrow 2\},S)$$.

[**定义2.2.1**] 称两个文法G_1与G_2**等价**,如果$L(G_1)=L(G_2)$.

[例2.2.4]

(1)语言$L=\{0^n\mid n\geq 1\}$的文法之一为:$G_1:S\to 0|0S$.

(2)语言$L=\{0^n\mid n\geq 0\}$的文法之一为:$G_2:S\to \varepsilon|0S$.

(3)语言$L=\{0^{2n}1^{3n}\mid n\geq 0\}$的文法之一为:$G_3:S\to \varepsilon|00S111$.

[例2.2.5] 构造一个文法,使得其语言$L(G)=\{w\mid w\in\{a,\cdots,z\}^+\}$.

[解] $G:S\to A|AS$,其中$A\to a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z$.

[例2.2.6] 构造一个文法,使得其语言$L(G)=\{ww^T\mid w\in\{0,1,2\}^+\}$.

[解] 直接构造文法较难想到,考虑递归定义L:

①对$\forall a\in\{0,1,2\}$,有$aa\in L$.

②若句子$x\in L$,则对$\forall a\in\{0,1,2\}$,有句子$axa\in L$.

③L中不包含不满足①和②的其他串.

根据递归定义可得产生式组:$G:S\to 00|11|22,S\to 0S0|1S1|2S2$.

[例2.2.6] 构造产生十进制有理数的文法.

[解] $G:S\to R|+R|-R,R\to N|B,B\to N.D,N\to 0|AM,D\to 0|MA,$

$$A\to 1|2|3|4|5|6|7|8|9,M\to\varepsilon|0M|1M|2M|3M|4M|5M|6M|7M|8M|9M.$$

[例2.2.7] 构造产生只包含加减乘除的算术表达式的文法,其中常数用c表示,变量用v表示.

[解] 考虑递归定义.

①基础:常数是算术表达式,变量是算术表达式.

②归纳:若E_1和E_2是算术表达式,则$E_1+E_2,E_1-E_2,E_1\times E_2,E_1/E_2$都是算术表达式.

③只有满足①和②的句子是算术表达式.

故$G:E\to c|v|E+E|E-E|E\times E|E/E$.

[例2.2.8] 构造一个语言$L(G)=\{a^nb^nc^n\mid n\geq 1\}$的文法.

[解] 注意$\{a^nb^nc^n\mid n\geq 1\}\neq\{a^nb^n\mid n\geq 1\}\{c^n\mid n\geq 1\}$.

注意到文法$G_1:S\to abc|aSbc$的语言$L(G_1)=\{a^n(bc)^n\mid n\geq 1\}$,可构造如下两种文法:

(1)$G_2:S\to aBC|aSBC,CB\to BC,aB\to ab,bB\to bb,bC\to bc,cC\to cc$.

(2)$G_3:S\to abc|aSBC,bB\to bb,cB\to Bc$.

[例2.2.9] 设字母表$\Sigma=\{a,b,c\}$.

(1)语言为$\{a^nb^n\mid n\geq 0\}$的文法之一为:$S\to aSb|\varepsilon$.

(2)语言为$\{a^nb^m \mid n,m\geq 1\}$的文法之一为:$S\rightarrow AB,A\rightarrow aA|a,B\rightarrow bB|b$.

[**例2.2.10**] 设字母表$\Sigma=\{a,b,c\}$.

(1)语言为$\{xwx^T \mid x,w\in\Sigma^+\}$的文法之一为:

$S\rightarrow aSa|bSb|cSc|aWa|bWb|cWc,W\rightarrow aW|bW|cW|a|b|c$.

(2)语言$\{xx^Tw \mid x,w\in\Sigma^+\}$的文法之一为:

$S\rightarrow XW,X\rightarrow aXa|bXb|cXc|aa|bb|cc,W\rightarrow aW|bW|bW|a|b|c$.

2.3 文法的Chomsky体系

[**定义2.3.1**] 文法的Chomsky体系

(1)任一文法$G=(V,T,P,S)$称为0**型文法**,又称**短语结构文法**(Phrase Structure Grammar, PSG),其对应的语言$L(G)$称为0**型语言**,又称**短语结构语言**(Phrase Structure Language, PSL)、**递归可枚举集**(recursively enumerable, r.e.).

(2)称文法$G=(V,T,P,S)$为1**型文法**或**上下文有关文法**(Context Sensitive Grammar, CSG),如果对$\forall(\alpha\rightarrow\beta)\in P$,都有$|\beta|\geq|\alpha|$.其对应的语言$L(G)$称为$1$**型语言**或**上下文有关语言**(Context Sensitive Language, CSL).

(3)称文法$G=(V,T,P,S)$为2**型文法**或**上下文无关文法**(Context Free Grammar, CFG),如果对$\forall(\alpha\rightarrow\beta)\in P$,都有$|\beta|\geq|\alpha|$且$\alpha\in V$.其对应的语言$L(G)$称为$2$**型语言**或**上下文无关语言**(Context Free Language, CFL).

(4)称文法$G=(V,T,P,S)$为3**型文法**或**正则文法**(Regular Grammar, RG),如果$\forall(\alpha\rightarrow\beta)\in P$都有形式$A\rightarrow w,A\rightarrow wB$,其中$A,B\in V,w\in T^+$.其对应的语言$L(G)$称为$3$**型语言**或**正则语言**(Regular Language, RL).

[**注**]

(1)①若文法G是RG,则它也是CFG、CSG、PSG,反之不然.若语言L是RL,则它也是CFL、CSL、PSL,反之不然.

②若文法G是CFG,则它也是CSG、PSG,反之不然.若语言L是CFL,则它也是CSL、PSL,反之不然.

③若文法G是CSG,则它也是PSG,反之不然.若语言L是CSG,则它也是PSL,反之不然.

(2)①文法G是CFG时,$L(G)$也可以是RL.

②文法G是CSG时,$L(G)$也可以是CFL、RL.

③文法G是PSG时,$L(G)$也可以是CSL、CFL、RL.

[**定理2.3.1**] 语言L是RL的充要条件是:$\exists$文法$G\ s.t.\ L=L(G)$,且其产生式要么形如$A\rightarrow a$,要么形如$A\rightarrow aB$,其中A和B为语法变量,a为终极符.

[**证**] (充) 若文法G的产生式要么形如$A\rightarrow a$,要么形如$A\rightarrow aB$,其中A和B为语法变量,a为终极符,

则它是RG,进而它产生的语言是RL.

(必) 用产生式组$A\rightarrow a_1A_1,A_1\rightarrow a_2A_2,\cdots,A_{n-1}\rightarrow a_n$代替产生式$A\rightarrow a_1\cdots a_n$,

用产生式组$A\rightarrow a_1A_1,A_1\rightarrow a_2A_2,\cdots,A_{n-1}\rightarrow a_nB$代替产生式$A\rightarrow a_1\cdots a_nB$.

设原文法为G,替换后的文法为G'.对推导步数归纳,可证明加强的结论:对$\forall A\in V$,有$(A\Rightarrow_G^+ x)\Leftrightarrow(A\Rightarrow_{G'}^+ x)$.

[**定义2.3.2**]

(1)对文法$G=(V,T,P,S)$,若$\forall(\alpha\rightarrow\beta)\in P$都有形式$A\rightarrow w$或$A\rightarrow wBx$,其中$A,B\in V,w,x\in T^*$,则称$G$为**线性文法**,其对应的语言$L(G)$称为**线性语言**.

(2)对文法$G=(V,T,P,S)$,若$\forall(\alpha\rightarrow\beta)\in P$都有形式$A\rightarrow w$或$A\rightarrow wB$,其中$A,B\in V,w\in T^*$,则称$G$为**右线性文法**,其对应的语言$L(G)$称为**右线性语言**.

(3)对文法$G=(V,T,P,S)$,若$\forall(\alpha\rightarrow\beta)\in P$都有形式$A\rightarrow w$或$A\rightarrow Bw$,其中$A,B\in V,w\in T^*$,则称$G$为**左线性文法**,其对应的语言$L(G)$称为**左线性语言**.

[**定理2.3.2**] 语言L是左线性语言的充要条件是:存在文法$G\ s.t.\ L(G)=L$,且G中的产生式都有形式$A\rightarrow a$或$A\rightarrow Ba$,其中$A,B\in V,a\in T$.

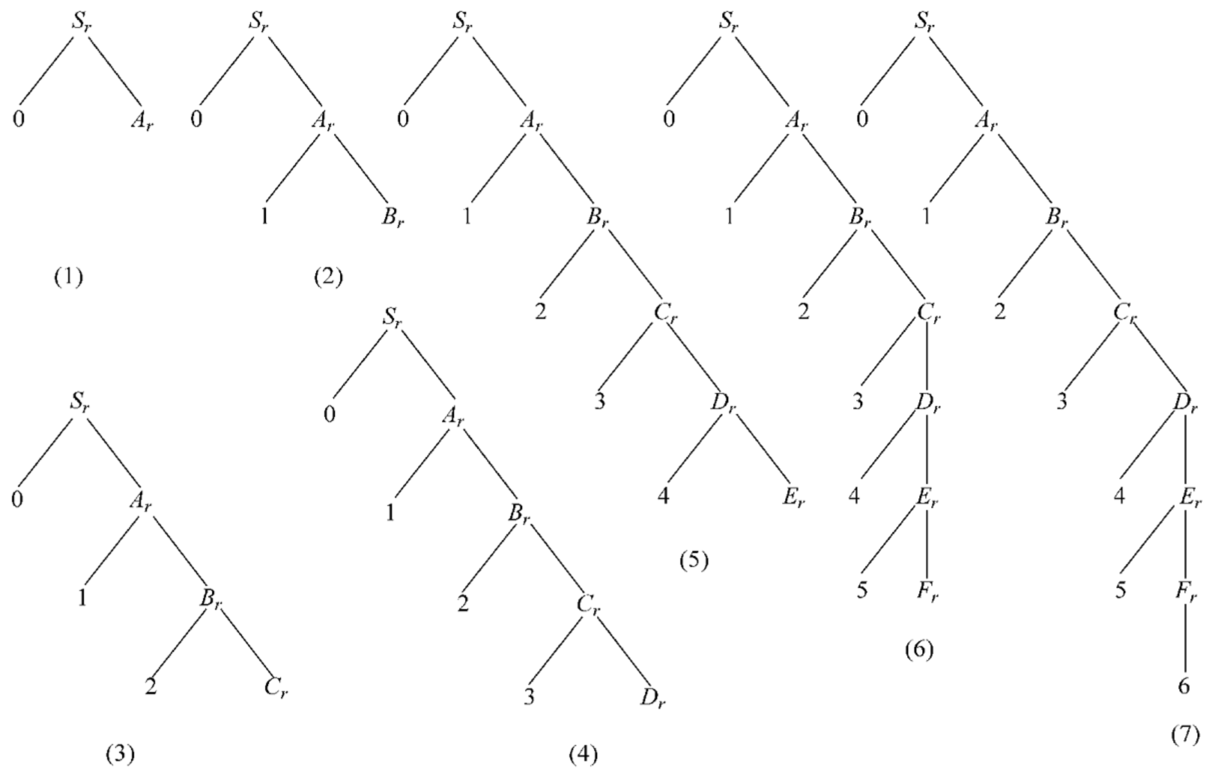
[**定理2.3.3**] 左线性文法于右线性文法等价.

[**例2.3.1**] 构造语言$\{0123456\}$的左线性文法、右线性文法.

[**解**]

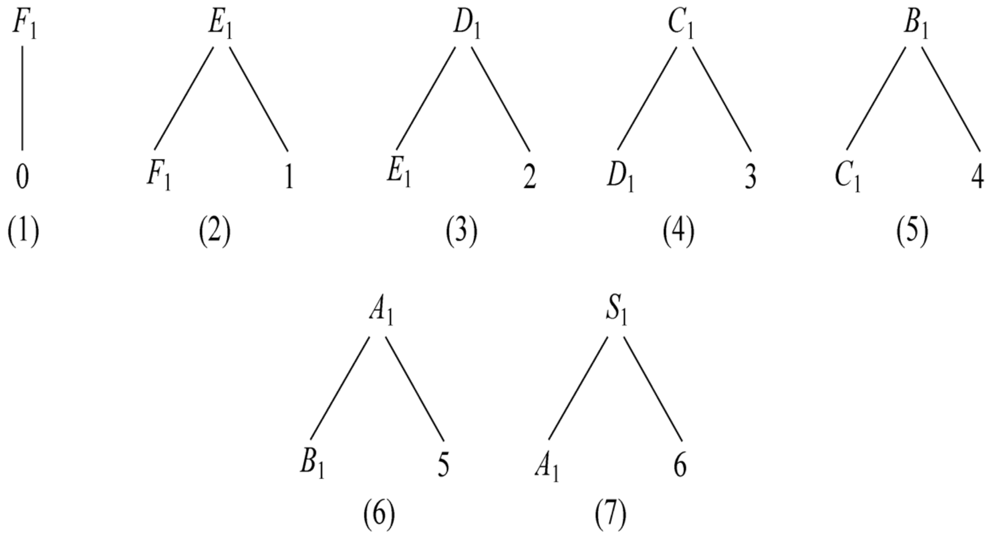
(1)右线性文法:$G_r:S_r\rightarrow 0A_r,A_r\rightarrow 1B_r,B_r\rightarrow 2C_r,C_r\rightarrow 3D_r,D_r=4E_r,E_r\rightarrow 5F_r,F_r\rightarrow 6.$

句子0123456在G_r中的推导可表示为下图:



(2)左线性文法:$G_l:S_l\rightarrow A_l6,A_l\rightarrow B_l5,B_l\rightarrow C_l4,C_l\rightarrow D_l3,D_l\rightarrow E_l2,E_l\rightarrow F_l1,F_l\rightarrow 0.$

句子0123456在G_l中的归约可表示为下图:



[定理2.3.4] 左线性文法、右线性文法的产生式混用得到的文法不是RG.

[解] 考察文法$G':S\to 0A,A\to S1|1$,其语言$L(G')=\{0^n1^n\mid n\geq 1\}$.

可以证明$L(G')$不是RG,进而$\nexists$正则文法$G\ s.t.\ L(G)=L(G')$.

[定义2.3.3] 形如$A\to\varepsilon$的产生式称为**空产生式**或ε**产生式**.

[注1] 因CSG中不含空产生式,进而CSL、CFL、RL中都不含空串ε.

[注2] 空串ε在语言中的存在不影响该语言的有穷描述的存在性,甚至除为生成空串外,空产生式可不用于该语言中其他句子的推导.

[注3] 允许CSL、CFL、RL中包含空串ε会更方便,因此允许CSG、CFG、RG中包含空产生式,允许CSL、CFL、RL中包含空串ε.

[定理2.3.5] 对任一文法$G=(V,T,P,S)$,$\exists$与G同类型的文法$G'=(V',T',P',S')\ s.t.\ L(G)=L(G')$,且$G'$中开始符号$S'$不出现在$G'$的产生式的右部.

[证] 只需证明G存在产生式使得开始符号S在该产生式的右部的情况.

考察文法$G'=(V\bigcup\{S'\},T,P',S')$,其中$P'=P\bigcup\{S'\to\alpha|(S\to\alpha)\in P\}$.

(1)下证$L(G')\subseteq L(G)$.

$\forall x\in L(G')$在G'中存在推导:

$S'\Rightarrow\alpha$ *使用P'中的产生式$S'\to\alpha$.

$\Rightarrow^*x$ *使用P'中除$S'\to\alpha$外的产生式.

即推导$\alpha\Rightarrow^*x$时使用的产生式都是P中的产生式,则推导$\alpha\Rightarrow^*x$在G中仍成立.

由P'的定义:$(S\to\alpha)\in P$,则上述x在G中存在推导:

$S\Rightarrow\alpha$ *使用P中的产生式$S\to\alpha$.

$\Rightarrow^*x$ *使用P中除$S\to\alpha$外的产生式.

故$L(G')\subseteq L(G)$.

(2)下证$L(G)\subseteq L(G')$.

$\forall x \in L(G)$在G中存在推导:

$S \Rightarrow \alpha$ *使用P中的产生式$S \to \alpha$.

$\Rightarrow^* x$ *使用P中除$S \to \alpha$外的产生式.

由P'的定义:上述x在G'中存在推导:

$S' \Rightarrow \alpha$ *使用P'中的产生式$S' \to \alpha$.

$\Rightarrow^* x$ *使用P'中除$S' \to \alpha$外的产生式.

故$L(G) \subseteq L(G')$.

[**定义2.3.4**] 对文法$G=(V,T,P,S)$,若S不出现在G的产生式的右部,则:

(1)若G是CSG,则也称$G'=(V,T,P\bigcup\{S \to \varepsilon\},S)$为CSG,其产生的语言$L(G')$也称为CSL.

(2)若G是CFG,则也称$G'=(V,T,P\bigcup\{S \to \varepsilon\},S)$为CFG,其产生的语言$L(G')$也称为CFL.

(3)若G是RG,则也称$G'=(V,T,P\bigcup\{S \to \varepsilon\},S)$为RG,其产生的语言$L(G')$也称为RL.

[**定理2.3.6**] 对语言L,有:

(1)若L是CSL,则语言$L'=L\bigcup\{\varepsilon\}$也是CSL.

(2)若L是CFL,则语言$L'=L\bigcup\{\varepsilon\}$也是CFL.

(3)若L是RL,则语言$L'=L\bigcup\{\varepsilon\}$也是RL.

[**证**] 以证明(1)为例.

设L是CSL,则$\exists$CSG $G=(V,T,P,S)\ s.t.\ L(G)=L$.不妨设$S$不出现在$G$的产生式的右部.

考察文法$G'=(V,T,P\bigcup\{S \to \varepsilon\},S)$.

因S不出现在G的产生式的右部,则产生式$(S \to \varepsilon)$不出现在任意非空串的推导中.

易证$L(G')=L(G)\bigcup\{\varepsilon\}$.因$G'$是CSG,则$L(G)\bigcup\{\varepsilon\}$是CSL.

[**定理2.3.7**] 对语言L,有:

(1)若L是CSL,则语言$L'=L\backslash\{\varepsilon\}$也是CSL.

(2)若L是CFL,则语言$L'=L\backslash\{\varepsilon\}$也是CFL.

(3)若L是RL,则语言$L'=L\backslash\{\varepsilon\}$也是RL.

[**证**] 以证明(1)为例.

设L是CSL,则$\exists$CSG $G=(V,T,P,S)\ s.t.\ L(G)=L$.

显然只需证明$\varepsilon \in L$的情况.不妨设S不出现在G的产生式的右部.

考察文法$G'=(V,T,P,S,\{S \to \varepsilon\},S)$.

因S不出现在G的产生式的右部,则产生式$(S \to \varepsilon)$不出现在$L(G)$的任意非空串的推导中.

易证$L(G')=L(G)\backslash\{\varepsilon\}$.因$G'$是CSG,则$L(G)\backslash\{\varepsilon\}$是CSL.