形式语言与自动机

李舟军 教授

北京航空航天大学计算机学院

lizj@buaa.edu.cn

第五章 下推自动机

5 下推自动机

本章将介绍形式语言理论中的另一种十分重要的机器系统,即下推自动机(Push Down Automata).

讨论下推自动机与上下文无关语言和上下文无关文法之间的关系,即要说明被下推自动机接收的语言类恰为上下文无关语言类.然后讨论上下文无关语言类的封闭性质.

证明程序设计语言不是真正的上下文无关语言,单字母表上的上下文无关语言都是正规集.

证明一类特殊的0型文法,即所谓的终端受限文法，它所产生的语言都是上下文无关语言，这就是$Baker$定理.

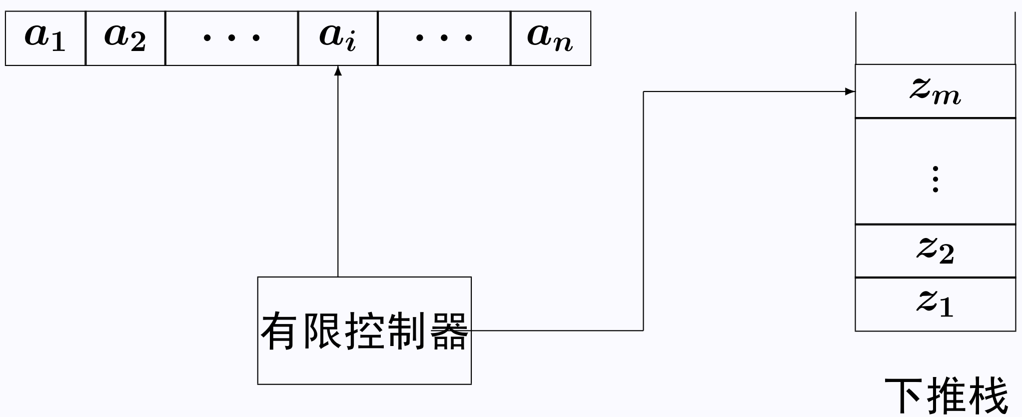
5.1. 下推自动机的形式定义

有穷自动机只能接收正规集,在对有穷自动机做各种改进之后(如允许非确定，ϵ动作和磁头的双向移动等)仍然如此，以致连十分简单的上下文无关文法语言$L = \{0^n1^n | n \geq 1\}$都不能接收,其原因主要是有穷自动机缺乏足够的记忆能力，虽然有穷自动机能用它的状态记忆信息，但因有穷自动机只有有限个状态，所以它能记忆的信息总量必有一个确切的上限，而要想接收语言$L = \{0^n1^n\}$就必须能识别出每个字$0^n1^n(n \geq 1)$,因此有穷自动机必须记住字0^n1^n中前面0的个数,因为n可以任意大,所以有穷自动机显然就无能为力了.

为了增强有穷自动机的接收能力,使它能够接收上下文无关语言,针对有穷自动机的不足，给它增加了一个容量无限的记忆装置–下推栈.下推栈就象一个长长的永远也装不满的口袋,每次装时都从口上装进去,把口袋里原来的东西都下压;每次取时也只能取最上面的靠近袋口的东西，所以，最先装进去的东西只能最后取出.

称如下图所示的带有一个下推栈的机械装置为下推自动机.

输入带



下推自动机M

下推自动机M由三部分组成:

1) 输入带

整个输入带被分割成有限个单元,每个单元都存放着一个输入符号,而且永远也不再改变，磁头正对着的单元，称为被扫描的带单元，并称这个带单元中所存放的输入符号为M的当前输入.如果输入带上各单元中所存放的输入符号,从左到右依次为$a_1, a_2, \cdots, a_n$时,就称字$w = a_1 a_2 \cdots, a_n$为对M的输入字或输入串,简称对M的输入;M的每次动作,都使输入磁头右移一个单元或者不动.

访问主页
标题页
第 7 页 共 88 页
返回
全屏显示
关闭
退出

2) 下推栈

下推栈是一个容量无限的先进后出存储器,它的第一个单元称为栈底,并称最后一个存放有信息的单元为栈顶,因此栈顶是随着栈内所存放的信息数量的改变而不断改变的.若从栈底到栈顶各单元所存放的信息依次为:$z_1, z_2, \cdots, z_m$时,则称字$\alpha = z_1 z_2 \cdots z_m$为栈内容.在开始时,栈内容为$z_0$,称$z_0$为起始栈符号.下推栈的磁头永远正对着栈顶单元, 因此,当栈内容为空字ϵ时,下推自动机M就停机.M的每次动作,下推栈内容的变化可以有以下四种情况:

1) 下推栈内容不变;

2) 仅栈顶符号改变;

3) 清除栈顶,新栈顶下移一个单元;

4) 清除下推栈顶符号,并重新向栈内输入$k+1$个符号,栈顶也上移k个单元($k \geq 1$);

3) 有限控制器

下推自动机M的有限控制器仅有有穷个内部状态,其中之一称为起始状态,常用q_0表示.还有若干个状态(也可能没有),称为M的终止状态或接收状态.M的每次动作,都使M的有限控制器从一种状态变到另一种状态.

下推自动机M的每次动作,都由它的当前输入,栈顶符号和M当前所处的状态(称为M的当前状态)所决定.

定义5.1.1 称七元偶$M =< K, \Sigma, \Gamma, \delta, q_0, Z_0, F >$为一个下推自动机,其中

K—-非空有穷集,称为M的状态集,并称每个$q \in K$为M的一个状态;

Σ—-非空有穷集,称为M的输入字母表,并称每个$a \in \Sigma$为M的一个输入符号;

Γ—-非空有穷集,称为M的栈符号集,并称每个$Z \in \Gamma$为M的一个栈符号;

δ—-为一个从$K \times \Sigma_\epsilon \times \Gamma$到$2^{K\times\Gamma^*}$的有穷函数,称为$M$的动作函数或变换函数;

q_0—-$q_0 \in K$,称为M的起始状态;

Z_0—-$Z_0 \in \Gamma$,称为M的起始栈符号,或栈底符号;

F—-$F \subseteq K$,称为M的终止状态集合(或接收状态集合),并称每个$q \in \Gamma$为M的一个终止状态(或接收状态);

下推自动机M也简记为$pda\ M$.

关于$pda\ M =< K, \Sigma, \Gamma, \delta, q_0, Z_0, F >$的几点说明

1) 变换函数δ为有穷函数;

2) 当$q \in K, a \in \Sigma_\epsilon, Z \in \Gamma$时,$\delta(q, a, Z) = \{< p_1, \gamma_1 >, \ldots, < p_t, \gamma_t >\}$的意义;

例1 取pda $M_1 =< \{q_0, q_1, q_2\}, \{0, 1\}, \{Z_0, A\}, \delta, q_0, Z_0, \{q_2\} >$,其中$\delta$定义如下:

i) $\delta(q_0, 0, Z_0) =< q_0, AZ_0 >;$

ii) $\delta(q_0, 0, A) =< q_0, AA >;$

iii) $\delta(q_0, 1, A) =< q_1, \epsilon >;$

iv) $\delta(q_1, 1, A) =< q_1, \epsilon >;$

v) $\delta(q_1, \epsilon, Z_0) =< q_2, \epsilon >;$

定义5.1.2 取pda $M =< K, \Sigma, \Gamma, \delta, q_0, z_0, F >$:

i) 若$q \in K, w \in \Sigma^*$且$\beta \in \Gamma^*$,则称$< q, w, \beta >$为M的一个瞬间描述或格局,简记为ID.

ii) 设$q \in K, a \in \Sigma_\epsilon, w \in \Sigma^*, Z \in \Gamma$且$\beta \in \Gamma^*$.若$< p, \gamma > \in \delta(q, a, Z)$,则称$M$可以从$< q, aw, Z\beta >$达到$< p, w, \gamma\beta >$, 记为$< q, aw, Z\beta > \vdash_M < p, w, \gamma\beta >$.

几点说明:

i) 在不致引起误解的情况下,常把$\vdash_M$简写为$\vdash$.

ii) 对pda $M =< K, \Sigma, \Gamma, \delta, q_0, Z_0, F >$,它所有的格局的集合显然为$K \times \Sigma^* \times \Gamma^*$,而$\vdash_M$为集合$K \times \Sigma^* \times \Gamma^*$上的一个二元关系.

iii) 对输入$w = a_1 \cdots a_n (a_1, \cdots, a_n) \in \Sigma$中的每个$a_i (1 \le i \le n)$,总认为在$a_i$的两边有任意多个$\epsilon$.因此,认为以下的输入$abb\epsilon\epsilon$ ϵabb $\epsilon\epsilon a\epsilon b\epsilon\epsilon b\epsilon\epsilon$都统统是输入$abb$.

iv) 格局$< q, w, \beta >$的意义,是指M处于状态q,输入带上的残缺输入为w,而当时的栈的内容为β.因此,M以后的动作就完全由$< q, w, \beta >$所决定.

例2: 对例1中的$pda\ M_1$及对它的输入字$w_1 = 0100101, w_2 = 001$和$w_3 = 0011$,显然有:

1) $< q_0, 010001, Z_0 > \vdash < q_0, 100101, AZ_0 > \vdash < q_1, 00101, Z_0 >$,这时$M_1$未用完对它的输入$w_1$而停机于非终止状态$q_1$.

2) $< q_0, 001, Z_0 > \vdash < q_0, 01, AZ_0 > \vdash < q_0, 1, AAZ_0 > \vdash < q_1, \epsilon, AZ_0 >$,这时$M_1$用完对它的输入$w_2$而停机于非终止状态$q_1$.

3) $< q_0, 0011, Z_0 > \vdash < q_0, 011, AZ_0 > \vdash < q_0, 11, AAZ_0 > \vdash < q_1, 1, AZ_0 > \vdash < q_1, \epsilon, Z_0 > \vdash < q_1, \epsilon, Z_0 > \vdash < q_2, \epsilon, \epsilon >$,这时$M_1$用完对它的输入$w_3$,下推栈变空并停机于接收状态$q_2$.

例3: 取pda $M_2 =< \{q_0, q_1, q_2\}, \{a, b\}, \{Z_0, A, B\}, \delta, q_0, Z_0, \{q_2\} >$,其中$\delta$定义如下:

$$
\begin{aligned}
&\delta(q_0, \epsilon, Z_0) =< q_2, \epsilon >;\\
&\delta(q_0, a, Z_0) = \{< q_2, AZ_0 >, < q_1, AZ_0 >\};\\
&\delta(q_0, a, A) = \{< q_2, AA >, < q_1, AA >\};\\
&\delta(q_0, b, Z_0) = \{< q_0, BZ_0 >, < q_1, BZ_0 >\};\\
&\delta(q_0, b, B) = \{< q_0, BB >, < q_1, BB >\};\\
&\delta(q_0, a, B) = \{< q_0, AB >, < q_1, AB >\};\\
&\delta(q_0, b, A) = \{< q_0, BA >, < q_1, BA >\};\\
&\delta(q_1, a, A) = \{< q_1, \epsilon >\};\\
&\delta(q_1, b, B) = \{< q_1, \epsilon >\};\\
&\delta(q_1, \epsilon, Z_0) = \{< q_2, \epsilon >, < q_2, Z_0 >\};\\
&\delta(q_2, \epsilon, Z_0) = \{< q_1, Z_0 >\};
\end{aligned}
$$

考虑M_2对输入$w = abbaabba$的动作:

1) $< q_0, abbaabba, Z_0 > \vdash < q_2, abbaabba, \epsilon >$

M_2未用一个输入符号就把下推栈变空,并停机于终止状态q_2.

2)

$$
\begin{aligned}
< q_0, abbaabba, Z_0 > &\vdash < q_0, bbaabba, AZ_0 > \\
&\vdash < q_0, baabba, BAZ_0 > \\
&\vdash < q_0, aabba, BBAZ_0 > \\
&\vdash < q_1, abba, ABBAZ_0 > \\
&\vdash < q_1, bba, BBAZ_0 > \\
&\vdash < q_1, ba, BAZ_0 > \\
&\vdash < q_1, a, AZ_0 > \\
&\vdash < q_1, \epsilon, Z_0 > \\
&\vdash < q_2, \epsilon, Z_0 > \\
&\vdash < q_1, \epsilon, Z_0 > \\
&\vdash < q_2, \epsilon, Z_0 > \\
&\vdash < q_1, \epsilon, Z_0 > \\
&\vdash < q_2, \epsilon, Z_0 >
\end{aligned}
$$

M_2既可以用完对它的输入w而把下推栈变空,并停机于终止状态q_2,也可以无限地做ϵ动作,时而进入终止状态q_2,时而又进入非终止状态q_1,循环往复,永不终止.

3)

$$
\begin{aligned}
< q_0, abbaabba, Z_0 > &\vdash\ < q_0, bbaabba, AZ_0 > \\
&\vdash\ < q_1, baabba, BAZ_0 > \\
&\vdash\ < q_1, aabba, AZ_0 > \\
&\vdash\ < q_1, abba, Z_0 >
\end{aligned}
$$

M_2未用完对它的输入而停机于非终止状态q_1.

pda从同一格局出发，可以到达不同的格局，而且pda的停机也有两种不同的方式:下推栈变空和$\delta(q,a,Z)$无定义.由此定义pda的两种不同的接收方式:

定义5.1.3 设$pda\ M=<K,\Sigma,\Gamma,\delta,q_0,Z_0,F>$且$w\in\Sigma^*$.

i) 用$\overset{+}{\vdash}_M$和$\overset{*}{\vdash}_M$分别表示$K\times\Sigma^*\times\Gamma^*$上的二元关系$\vdash_M$的传递闭包和自反传递闭包.

ii) 若有$q\in F$及$\beta\in\Gamma^*$,使$<q_0,w,Z_0>\overset{*}{\vdash}_M<q,\epsilon,\beta>$,则称$M$以终止状态(或接收状态)接收输入$w$.

iii) 若有$q\in K$使$<q_0,w,Z_0>\overset{*}{\vdash}_M<q,\epsilon,\epsilon>$,则称$M$以空栈接收输入$w$.

iv) 令:

$$T(M)=\{w\in\Sigma^*|\text{有}q\in F\text{及}\beta\in\Gamma^*\text{使}<q_0,w,Z_0>\overset{*}{\vdash}_M<q,\epsilon,\beta>\}$$

$$N(M)=\{w\in\Sigma^*|\text{有}q\in K\text{使}<q_0,w,Z_0>\overset{*}{\vdash}_M<q,\epsilon,\epsilon>\}$$

称$T(M)$为M以终止状态(或接收状态)接收的语言,而$N(M)$则称为M以空栈接收的语言.

对例1和例3中的$pda\ M_1$和$pda\ M_2$进行分析,可知有:

$$T(M_1) = N(M_1) = \{0^n1^n|n \geq 1\};$$

$$T(M_2) = N(M_2) = \{ww^R|w \in \{a,b\}^*\};$$

并非对所有的$pda\ M$都能有$T(M) = N(M)$.

例4: 取$M_3 =< \{q_0, q_1\}, \{0, 1, *\}, \{Z_0, Z, A, B\}, \delta, q_0, Z_0, \{q_1\} >$,其中$\delta$定义如下:

i) 接收输入$w = *$

$$\delta(q_0, *, Z_0) =< q_1, Z >$$

ii) 把输入$w = x * y$中的x输入栈中

$$\delta(q_0, 0, Z_0) =< q_0, AZ > \qquad \delta(q_0, 1, Z_0) =< q_0, BZ_0 >$$
$$\delta(q_0, 0, A) =< q_0, AA > \qquad \delta(q_0, 1, A) =< q_0, BA >$$
$$\delta(q_0, 0, B) =< q_0, AB > \qquad \delta(q_0, 1, B) =< q_0, BB >$$

iii) 比较x^R和y

$$\delta(q_0, *, A) =< q_1, A > \qquad \delta(q_0, *, B) =< q_1, B >$$
$$\delta(q_1, 0, A) =< q_1, \epsilon > \qquad \delta(q_1, 1, B) =< q_1, \epsilon >$$
$$\delta(q_1, \epsilon, Z_0) =< q_0, \epsilon >$$

显然M_3为pda,而且不难用归纳法证明

$$T(M_3) = \{vu * u^R | u, v \in \{0, 1\}^*\}$$
$$N(M_3) = \{1w * w^R 1 | w \in \{0, 1\}^*\}$$

因此,$T(M_3) \neq N(M_3)$.

定理5.1.1 设$pda\ \ M = < K, \Sigma, \Gamma, \delta, q_0, Z_0, F >, q \in K, x, y, w \in \Sigma^*, \alpha, \beta, \gamma \in \Gamma^*$且$n, m \in \mathbb{N}$,如果

$$< q, x, \alpha > \overset{n}{\vdash} < p, \epsilon, \epsilon >$$

且

$$< p, yw, \beta\gamma > \overset{m}{\vdash} < \bar{q}, w, \gamma >$$

则

$$< q, xyw, \alpha\beta\gamma > \overset{n+m}{\vdash} < \bar{q}, w, \gamma >.$$

访 问 主 页
标 题 页
第 23 页 共 88 页
返 回
全 屏 显 示
关 闭
退 出

证明:用关于n的归纳法来证.

当$n=0$时,则$<q,x,\alpha>=<p,\epsilon,\epsilon>$.此时命题为真.

假定对任意的自然数$k\geq 0$,当$n=k$时命题为真.

当$n=k+1$时,$<q,x,\alpha>\overset{k+1}{\vdash}<p,\epsilon,\epsilon>$必有如下形式:

$$<q,ax',Z\alpha'>\vdash<q',x',\gamma'\alpha'>\overset{k}{\vdash}<p,\epsilon,\epsilon>$$

其中$x=ax',\alpha=Z\alpha',a\in\Sigma_\epsilon,Z\in\Gamma$且$<q',\gamma'>\in\delta(q,a,Z)$.从而由归纳假设得

$<q',x'yw,\gamma'\alpha'\beta\gamma>\overset{k+m}{\vdash}<\bar{q},w,\gamma>$,因此, 再由$<q',\gamma'>\in\delta(q,a,Z),x=ax'$和$\alpha=Z\alpha'$知道

$$<q,xyw,\alpha\beta\gamma>\overset{k+1+m}{\vdash}<\bar{q},w,\gamma>.$$

定理5.1.2 设pda $M =< K, \Sigma, \Gamma, \delta, q_0, Z_0, F >, q, q_1 \in K, w \in \Sigma^*$且$B_1, \cdots, B_s \in \Gamma$,若$x \in \Sigma^*$和$\beta \in \Gamma^*$使

$$< q_1, wx, B_1 \cdots B_s\beta > \overset{n}{\vdash} < q, x, \beta >$$

则必有$u_1, \cdots, u_s \in \Sigma^*, q_2, \cdots, q_s \in K$及$n_1, \cdots, n_s \in \mathbb{N}$,使

$$w = u_1 \cdots u_s \qquad n = n_1 + \cdots + n_s$$

且

$$< q_i, u_i, B_i > \overset{n_i}{\vdash} < q_{i+1}, \epsilon, \epsilon > \quad (1 \le i \le s)$$

其中$q_{s+1} = q$.

证明:用关于n的第二归纳法来证明.

当$n = 0$时,则$< q_1, wx, B_1 \ldots B_s\beta > = < q, x, \beta >$,所以$q_1 = q, w = \epsilon$且$B_1 \ldots B_s = \epsilon$,显然,此时命题为真.

假定对任意正整数k,当$n < k$时命题为真.

当$n = k \geq 1$时,$< q_1, wx, B_1 \ldots B_s\beta > \overset{k}{\vdash} < q, x, \beta >$必有如下形式:

$< q_1, aw_1x, B_1 \ldots B_s\beta > \vdash < q_1', w_1x, A_1 \ldots A_tB_2 \ldots B_s\beta > \overset{k-1}{\vdash} < q, x, \beta >$

其中$w = aw_1$,且$< q_1', A_1 \ldots A_t > \in \delta(q_1, a, B_1)$.根据归纳假设,必有$v_1, \ldots, v_t, u_2, \ldots, u_s \in \Sigma^*, q_2', \ldots, q_t', q_2, \ldots, q_s \in K$以及$l_1, \ldots, l_t, k_1, \ldots, k_s \in \mathbb{N}$,使

$w_1 = v_1 \ldots v_tu_2 \ldots u_s, k - 1 = l_1 + \cdots + l_t + k_2 + \cdots + k_s$

且$< q_j', v_j, A_j > \overset{l_j}{\vdash} < q_{j+1}', \epsilon, \epsilon > (1 \leq j \leq t)$

$< q_i, u_i, \beta_i > \overset{k_i}{\vdash} < q_{i+1}, \epsilon, \epsilon > (2 \leq i \leq s)$

其中$q_{t+1}' = q_2$.

访问主页
标题页
◀◀ ▶▶
◀ ▶
第 26 页 共 88 页
返回
全屏显示
关闭
退出

从而由定理5.1.1知道

$$
\begin{aligned}
< q_1, av_1 \dots v_t, B_1 > &\vdash\ < q_1', v_1 \dots v_t, A_1 \dots A_t > \\
&\overset{l_1}{\vdash}\ < q_2', v_2 \dots v_t, A_2 \dots A_t > \\
&\ \vdots \\
&\overset{l_{t-1}}{\vdash}\ < q_t', v_t, A_t > \\
&\overset{l_t}{\vdash}\ < q_2, \epsilon, \epsilon >
\end{aligned}
$$

因此,当取$u_1 = av_1 \dots v_t$及$k_1 = 1 + l_1 + \cdots + l_t$时,由

$$< q_1, u_1, B_1 > \overset{k_1}{\vdash} < q_2, \epsilon, \epsilon >$$

即知,当$n = k$时,命题也真.

5.2. 上下文无关语言和下推自动机

定理5.2.1 设Σ为字母表,且$L \subseteq \Sigma^*$,则以下五个条件互相等价:

i) L为cfl.

ii) 有单状态的pda M_1,使$N(M_1) = L$.

iii) 有两状态的pda M_2,使$T(M_2) = L$.

iv) 有pda M_3,使$T(M_3) = L$.

v) 有pda M_4,使$N(M_4) = L$.

证明:

$i) \Rightarrow ii)$

因为L为cfl,故有$Greibach$范式文法$G =< V_N, \Sigma, P, S >$,使$L(G) = L \setminus \{\epsilon\}$,令:

$M_1 =< \{q_0\}, \Sigma, V_N, \delta, q_0, S, \emptyset >$

其中δ定义如下:

$\delta(q_0, \epsilon, S) = \{< q_0, \epsilon > |\epsilon \in L\}$

$\delta(q_0, a, A) = \{< q_0, \gamma > |\gamma \in V_N^*$且$A \rightarrow a\gamma \in P\}, A \in V_N$且$a \in \Sigma$

显然M_1是一个pda.

若$A \in V_N, x \in \Sigma^*$且$\alpha, \beta \in v_N^*$,则有:

$xA\beta \Rightarrow xa\alpha\beta$当且仅当$< q_0, a, A\beta > \vdash < q_0, \epsilon, \alpha\beta >$

因此,由定理5.1.1得:

$S \overset{*}{\Rightarrow} w \in \Sigma^*$当且仅当$< q_0, w, S > \overset{*}{\vdash} < q_0, \epsilon, \epsilon >$

这表明$N(M_1) = L(G)$,即$N(M_1) = L$.

$ii) \Rightarrow iii)$

任取单状态$pda\ M_1 =< \{q_0\}, \Sigma, \Gamma, \delta, q_0, Z_0, \emptyset >$,令

$M_2 =< \{q_0, q_f\}, \Sigma, \Gamma \cup \{X_0, X\}, \delta', q_0, X_0, \{q_f\} >$

其中$X_0, X \notin \Gamma$为任意两个不同的新符号,$q_0 \neq q_f$,而δ'定义如下:

$\delta'(q_0, \epsilon, X_0) =< q_0, Z_0X >$

$\delta'(q_0, a, Z) = \delta(q_0, a, Z)$

$\delta'(q_0, \epsilon, X) =< q_f, \epsilon >$

这里$a \in \Sigma_\epsilon$且$Z \in \Gamma$.

显然M_1为一个两状态下推自动机,而且还有

$< q_0, w, Z_0 > \overset{*}{\vdash}_{M_1} < q_0, \epsilon, \epsilon >$

当且仅当$< q_0, w, Z_0X > \overset{*}{\vdash}_{M_2} < q_0, \epsilon, X >$

因此$< q_0, w, Z_0 > \overset{*}{\vdash}_{M_1} < q_0, \epsilon, \epsilon >$

当且仅当$< q_0, w, X_0 > \overset{*}{\vdash}_{M_2} < q_f, \epsilon, \epsilon >$

从而得到$w \in N(M_1)$当且仅当$w \in T(M_2)$.

这表明$T(M_2) = N(M_1)$.

$iii) \Rightarrow iv)$

显然成立.

$iv) \Rightarrow v)$

对任意的$pda\ M_3 =< K, \Sigma, \Gamma, \delta, q_0, Z_0, F >$,任取$q_f \notin K$及两个不同的新符号$X_0, X \notin \Gamma$,令

$M_4 =< K \cup \{q_f\}, \Sigma, \Gamma \cup \{X_0, X\}, \delta', q_0, X_0, \emptyset >$

其中δ'定义如下:

(1)开始

$\delta'(q_0, \epsilon, X_0) =< q_0, Z_0X >$

(2)模拟

若$q \in K, a \in \Sigma$,且$Z \in \Gamma$,则令

$\delta'(q, a, Z) = \delta(q, a, Z)$

$\delta'(q, \epsilon, Z) = \delta(q, \epsilon, Z) \cup \{< q_f, \epsilon > | q \in F\}$

$\delta'(q, \epsilon, X) = \{< q_f, \epsilon > | q \in F\}$

(3)接收

$\delta'(q_f, \epsilon, Z) =< q_f, \epsilon > \qquad (Z \in \Gamma)$

$\delta'(q_f, \epsilon, X) =< q_f, \epsilon >$

显然M_4为一个pda,而且还有

$< q_0, w, Z_0 > \overset{*}{\vdash}_{M_3} < q, \epsilon, \beta >$

当且仅当$< q_0, w, Z_0X > \overset{*}{\vdash}_{M_4} < q, \epsilon, \beta X >$

因此$< q_0, w, X_0 > \overset{*}{\vdash}_{M_4} < q_f, \epsilon, \epsilon >$

当且仅当有$q \in F$及$\beta \in \Gamma^*$使$< q_0, w, Z_0 > \epsilon \overset{*}{\vdash}_{M_3} < q, \epsilon, \beta >$

从而得到$w \in N(M_4)$当且仅当$w \in T(M_3)$.

这表明$N(M_4) = T(M_3)$.

$v) \Rightarrow i)$

对任意的$pda\ M =< K, \Sigma, \Gamma, \delta, q_0, Z_0, \emptyset >$,令

$G_M =< V_N, \Sigma, P, S >$

其中$V_N = \{S\} \cup \{[q, A, p] | q, p \in K$且$A \in T\}$

而P由以下三组产生式组成:

第一组

$S \rightarrow [q_0, Z_0, q], q \in K$

第二组

$[q, A, p] \rightarrow a$,若$< p, \epsilon > \in \delta(q, a, A)$

第三组

$[q, A, p] \quad \rightarrow \quad a[q_1, B_1, q_2] \ldots [q_{m-1}, B_{m-1}, q_m][q_m, B_m, p]$,若$< q_1, B_1 \ldots B_m > \in \delta(q, a, A)$且$q_2, \ldots, q_m \in K$

显然G_M为一个广义的cfg.

关于$L(G_M) = N(M)$的证明,分以下两步来完成:

1) 设$[q, A, p] \in V_N$,且$w \in \Sigma^*$.若$< q, w, A > \overset{*}{\vdash} < p, \epsilon, \epsilon >$,则$[q, A, p] \overset{*}{\Rightarrow} w$.

因为$< q, w, A > \overset{*}{\vdash} < p, \epsilon, \epsilon >$,所以有$n \in \mathbb{N}$,使$< q, w, A > \overset{n}{\vdash} < p, \epsilon, \epsilon >$.以下就用关于$n$的第二归纳法来证明.

首先,当$n = 1$时,命题显然为真;

其次,假定对任意的正整数$k > 1$,当$n < k$时命题为真.

当$n = k$时,由于$k \geq 1$,所以$< q, w, A > \overset{k}{\vdash} < p, \epsilon, \epsilon >$必有如下形式:

$$< q, aw', A > \vdash < q_1, w', B_1 \ldots B_l > \overset{k-1}{\vdash} < p, \epsilon, \epsilon >$$

其中$w = aw'$且$< q_1, B_1 \ldots B_l > \in \delta(q, a, A)$.根据定理5.1.2,这时有$x_1, \ldots, x_l \in \Sigma^*$及$q_2, \ldots, q_l \in K$,使$w' = x_1 \ldots x_l$

且$< q_i, x_i, B_i > \overset{k_i}{\vdash} < q_{i+1}, \epsilon, \epsilon > \ (1 \leq i \leq l)$

因此

$$\begin{aligned}
[q, A, p] &\Rightarrow a[q_1, B_1, q_2] \ldots [q_l, B_l, q_{l+1}] \\
&\overset{*}{\Rightarrow} ax_1[q_2, B_2, q_3] \ldots [q_l, B_l, q_{l+1}] \\
&\vdots \\
&\overset{*}{\Rightarrow} ax_1 \ldots x_{l-1}[q_l, B_l, q_{l+1}] \\
&\overset{*}{\Rightarrow} ax_1 \ldots x_{l-1}x_l
\end{aligned}$$

亦即$[q, A, p] \overset{*}{\Rightarrow} w$.

这表明当$n = k$时命题为真.

2) 设$[q,A,p] \in V_N$,且$w \in \Sigma^*$.若$[q,A,p] \overset{*}{\Rightarrow} w$,则$< q,w,A > \overset{*}{\vdash} < p,\epsilon,\epsilon >$.

对此,用关于派生$[q,A,p] \overset{*}{\Rightarrow} w$中所用派生步数$n$的第二归纳法来证明.

首先,当$n=1$时,命题显然真.

其次,假定对任意的正整数$k>1$,当$n<k$时命题为真.

当$n=k$时,由于$k>1$,所以$[q,A,p] \overset{k}{\Rightarrow} w$必有如下形式:

$[q,a,p] \Rightarrow a[q_1,B_1,q_2]\ldots[q_{l-1},B_{l-1},q_l][q_l,B_l,p] \overset{k-1}{\Rightarrow} w$

因为派生的第一步只能用第三组中的产生式.当记$q_{l+1}=p$时,根据引理2.2,必有$x_0=a$及$x_1,\ldots,x_l \in \Sigma^*$,使$w=x_0x_1\ldots x_l$

且$[q_i,B_i,q_{i+1}] \overset{k_i}{\Rightarrow} x_i \quad 1 \le i \le l$

其中$k-1=k_1+\cdots+k_l$

从而由$k_i \le k-1(1 \le i \le l)$及归纳假设可得

$< q_i,x_i,B_i > \overset{*}{\vdash} < q_{i+1},\epsilon,\epsilon >$

所以由定理5.1.1得到

$< q_1, x_0x_1 \ldots x_l, A > \overset{*}{\vdash} < p, \epsilon, \epsilon >$即$< q, w, A > \overset{*}{\vdash} < p, \epsilon, \epsilon >$

这表明,当$n = k$时命题为真.

由1)和2)可知

$w \in L(G_M)$当且仅当$w \in N(M)$

所以$L(G_M) = N(M)$.

例1 设$pda\ M =< \{q_0, q_1\}, \{0, 1\}, \{Z_0, X\}, \delta, q_0, Z_0, \emptyset >$,其中$\delta$定义如下:

$$\delta(q_0, 0, Z_0) =< q_0, XZ_0 >$$
$$\delta(q_0, 0, X) =< q_0, XX >$$
$$\delta(q_0, 1, X) =< q_1, \epsilon >$$
$$\delta(q_1, \epsilon, Z_0) =< q_1, \epsilon >$$
$$\delta(q_1, \epsilon, X) =< q_1, \epsilon >$$
$$\delta(q_1, 1, X) =< q_1, \epsilon >$$

试构造一个$cfg\ G$,使$L(G) = N(M)$.

例2 对例1中所得到的$cfg\ \tilde{G}$,试构造一个$pda\ \tilde{M}$, 使$T(\tilde{M}) = L(\tilde{G})$.

5.3. 上下文无关语言的性质

定理5.3.1 cfl类关于运算$\cup$(并),$\cdot$(连接), R(逆)和$*$闭包都封闭.

证明：设L和L'为任意两个cfl,则有$cfg\ G =< V_N, V_T, P, S >$和$G' =< V'_N, V'_T, P', S' >$,使$L(G) = L$且$L(G') = L'$.不失一般性,还可以假定

$$V_N \cup (V'_N \cup V'_T) = V'_N \cup (V_N \cup V_T) = \emptyset$$

任取$\tilde{S} \notin V_N \cup V'_N \cup V_T \cup V'_T$,并令

$$G_1 =< V_N \cup V'_N \cup \{\tilde{S}\}, V_T \cup V'_T, P \cup P' \cup \{\tilde{S} \to S, \tilde{S} \to S'\}, \tilde{S} >$$

$$G_2 =< V_N \cup V'_N \cup \{\tilde{S}\}, V_T \cup V'_T, P \cup P' \cup \{\tilde{S} \to SS'\}, \tilde{S} >$$

$$G_3 =< V_N, V_T, \{A \to \alpha^R | A \to \alpha \in P\}, S >$$

$$G_4 =< V_N \cup \{\tilde{S}\}, V_T, P \cup \{\tilde{S} \to S\tilde{S}, \tilde{S} \to \epsilon\}, \tilde{S} >$$

则G_1, G_2, G_3, G_4都是cfg,且有

$$L(G_1) = L \cup L'$$

$$L(G_2) = L \cdot L'$$

$$L(G_3) = L^R$$

$$L(G_4) = L^*$$

所以$L_1 \cup L', L \ldots L', L^R$和$L^*$都是$cfl$.

访问主页
标题页
第 41 页 共 88 页
返回
全屏显示
关闭
退出

定理5.3.2 cfl类关于置换封闭.

证明： 设$\Sigma = \{a_1, \ldots, a_n\}$,且$L \subseteq \Sigma^*$为$cfl$.若置换$f : \Sigma \rightarrow 2^{\Delta^*}$,使$f(a_1), \ldots, f(a_n)$都是$cfl$,则必有$Greibach$范式文法

$G_i =< V_N^{(i)}, V_T^{(i)}, P^{(i)}, S^{(i)} > \quad i = 0, \ldots, n$

使$L(G_0) = L$且$L(G_j) = f(a_j) \quad j = 1, 2, \ldots, n$

不失一般性,假定$V_N^{(i)} \cap (V_N^{(j)} \cup V_T^{(j)}) = \emptyset (i \neq j)$.若令

$G =< \cup_{i=0}^{n} V_N^{(i)}, \cup_{i=0}^{n} V_T^{(i)}, P, S^(0) >$

其中,$P = \{A \rightarrow S^{(i)}\alpha | A \rightarrow a_i\alpha \in P^{(0)}\} \cup P^{(1)} \cup \cdots \cup P^{(n)}$

则G为cfg,而且还有

$S^{(0)} \overset{*}{\Rightarrow}_{G_0} \quad L \quad a_{i_1}a_{i_2}\ldots a_{i_m}$当且仅当$S^{(0)} \overset{*}{\Rightarrow}_{G} \quad L \quad f(a_{i_1})f(a_{i_2})\ldots f(a_{i_m})$

所以$L(G) = f(L(G_0)) = f(L)$,即$f(L)$为cfl.

访问主页
标题页
◀◀ ▶▶
◀ ▶
第 42 页 共 88 页
返回
全屏显示
关闭
退出

定理5.3.3 cfl类关于同态封闭.

因为同态也是一种置换,所以定理5.3.3只是定理5.3.2的一种特殊情况.

定理5.3.4 cfl类关于逆同态封闭.

证明:设$L \subseteq \Delta^*$为cfl,则有pda $M =< K, \Delta, \Gamma, \delta, q_0, Z_0, F >$,使$T(M) = L$.对任意的同态$h : \Sigma \rightarrow \Delta^*$,当记$l = max|h(a)|a \in \Sigma$时,令

$\bar{M} =< \bar{K}, \Sigma, \Gamma, \delta, [q_0, \epsilon], z_0, \bar{F} >$

其中$\bar{K} = \{[q, x]|q \in K, x \in \Delta^*$且$|x| \leq l\}$

$\bar{F} = \{[q, \epsilon]|q \in F\}$

而δ'定义如下:

1) $\bar{M}$读输入符号

若$q \in K$,$a \in \Sigma$且$Y \in \Gamma$,则令

$\delta'([q, \epsilon], a, Y) =< [q, h(a)], Y >$

这表明,$\bar{M}$只是把当前的输入a变为$h(a)$而暂时存在它自己的有限控制器里面.

2) $\bar{M}$模拟M对输入符号串$h(a)$的动作

若$b \in \Delta_\epsilon, x \in \Delta^*, Y \in \Gamma, q \in K$且$|bx| \leq l$,则令

$\delta'([q, bx], \epsilon, Y) = \{[p, x], \gamma | < p, \gamma > \in \delta(q, b, Y)\}$

这表明,$\bar{M}$按其有限控制器里存放的$h(a)$动作,直到用完为止.

显然$\bar{M}$为一个pda.下面来证明$h^{-1}(T(M)) = T(\bar{(}M))$.

对任意的$a \in \Sigma$,因为

$< q, h(a), \alpha > \overset{*}{\vdash}_M < p, \epsilon, \beta >$

当且仅当$< [q, \epsilon], a, \alpha > \vdash_M < [q, h(a)], \epsilon, \alpha > \overset{*}{\vdash}_M < [p, \epsilon], \epsilon, \beta >$

从而不难由按字长的归纳法知道,若$w \in \Sigma^*$,则

$< [q, \epsilon], w, \alpha > \overset{*}{\vdash}_M < [p, \epsilon], \epsilon, \beta >$

当且仅当$< q, h(w), \alpha > \overset{*}{\vdash}_M < p, \epsilon, \beta >$

因此有$w \in T(\bar{M})$当且仅当$h(w) \in T(M)$即$w \in h^{-1}(T(M))$.

所以$T(\bar{(}M)) = h^{-1}(T(M)) = h^{-1}(L)$,即$h^{-1}(L)$为$cfl$.

定理5.3.5 cfl类关于运算$\cap$和$\sim$ (求补)都不封闭.

证明： 取$L_1 = \{a^i b^i c^j | i, j \geq 1\}$和$L_2 = \{a^i b^i c^j | i, j \geq 1\}$

显然L_1和L_2都是cfl.但因

$\sim (\sim L_1 \cup \sim L_2) = L_1 \cap L_2 = \{a^i b^i c^i | i \geq 1\}$

由于$\{a^i b^i c^i | i \geq 1\}$不是$cfl$,所以$cfl$类关于运算$\cup$和$\sim$都不封闭.

定理5.3.6 cfl类关于运算$\cap_R$(与正规集之交)封闭.

证明: 任取$cfl\ L \subseteq \Sigma^*$及正规集$R \subseteq \Sigma^*$.

则有$pda\ P =< K_P, \Sigma, \Gamma, \delta_P, Z_0, p_0, F_P >$,使$T(P) = L$.

还有$\Sigma_{dfa} A =< K_A, \Sigma, \delta_A, q_0, F_A >$,使$T(A) = R$.令

$M =< K_P \times K_A, \Sigma, \Gamma, \delta, [p_0, q_0], Z_0, F_P \times F_A >$

其中δ定义如下:若$[p, q] \in K_P \times K_A, a \in \Sigma_\epsilon$且$Z \in \Gamma$,则令

$\delta([p, q], a, Z) = \{< [p', \delta_A(q, a)], \gamma > \mid < p', \gamma > \in \delta_P(p, a, Z)\}$

显然M为一个pda

因为$w \in T(P) \cap T(A)$,当且仅当有$a_1, \ldots, a_n \in \Sigma_\epsilon, p_0, \ldots, p_n \in K_P, q_0, \ldots, q_n \in K_A$及$\gamma_1, \ldots, \gamma_n \in \Gamma^*$,使$w = a_1 \ldots a_n, p_n \in F_P, q_n \in F_A$,且$\delta_A(q_i, a_{i+1}) = q_{i+1}$. $< p_i, a_{i+1}, \gamma_i > \vdash_P < p_{i+1}, \epsilon, \gamma_{i+1} > (0 \le i < n)$,其中$\gamma_0 = Z_0$.

亦即$< [p_i, q_i], a_{i+1}, \gamma_i > \vdash_M < [p_{i+1}, q_{i+1}], \epsilon, \gamma_{i+1} > (0 \le i < n)$

所以$w \in T(P) \cap T(A)$当且仅当$w \in T(M)$

这表明$L \cap R = T(P) \cap T(A) = T(M)$为上下文无关语言.

访问主页
标题页
第 47 页 共 88 页
返回
全屏显示
关闭
退出

定理5.3.7 cfl类关于运算$/_R$(与正规集之商)封闭.

证明： 设$L \subseteq \Sigma^*$为cfl且$R \subseteq \Sigma^*$为rl.当记$\Sigma' = \{a' | a \in \Sigma\}$时,定义同态$h_1 : \Sigma \cup \Sigma' \rightarrow \Sigma^*$和$h_2 : K_P \times K_A \Sigma \cup \Sigma' \rightarrow \Sigma^*$如下:

$$h_1(a) = h_1(a') = a$$

$$h_2(a) = \epsilon, h_2(a') = a$$

这时显然有$L \setminus R = h_2(h_1^{-1}(L) \cap (\Sigma')^* \cdot R)$

根据定理5.3.4和推论5.3.3知道,L/R为上下文无关语言.

例1 取

$$L_1 = \{a^i b^i | i \geq 1\}^* a$$
$$L_2 = \{b^{2i} a^i | i \geq 1\}^*$$

则L_1和L_2都是cfl,但是L_2/L_1不是cfl.

5.4. 广义时序机变换

定义5.4.1 称六元偶$M =< K, \Sigma, \Delta, \delta, q_0, F >$为广义时序机(简记为$gsm\ M$),其中:

i) K, Σ, Δ, q_0和F的意义与时序机相同.

ii) δ为一个从$K \times \Sigma$到$2^{K \times \Delta^*}$的有穷函数,即对任意的$q \in K$及$a \in \Sigma$,皆有$\#\delta(q, a) < \infty$

称为M的动作函数.

若对任意的$p, q \in K$及$a \in \Sigma$,皆有$< p, \epsilon > \notin \delta(q, a)$,则称$gsm\ M$为$\epsilon$—无关的.

例1 令

$$M = < \{q_0, q_1\}, \{0, 1\}, \{a, b\}, \delta, q_0, \{q_1\} >$$

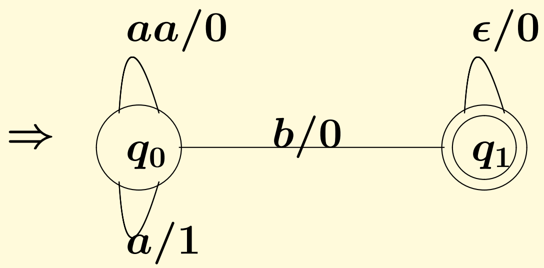
其中δ定义如下:

	0	1
q_0	$\{< q_0, aa >, < q_1, b >\}$	$< q_0, a >$
q_1	$\backslash$	$< q_1, \epsilon >$

则M是一个gsm,但不是$\epsilon-$无关的.

gsm也可以用状态图表示.它与Σ_{fa}的状态图完全类似，差异仅仅是弧上标记的符号由原来的输入符号a变为现在的w/a,其中a表示当前的输入,w表示当前的输出.

对例1的$gsm\ M$,其状态图如下图所示：



对$gsm\ M =< K,\Sigma,\Delta,\delta,q_0,F >$,也要把函数$\delta$加以延拓,使它成为一个从$K\times\Sigma^*$到$2^{K\times\Delta^*}$的函数$\widehat{\delta}$.为此,对任意的$q\in K, a\in\Sigma$及$x\in\Sigma^*$,令

$$\widehat{\delta}(q,\epsilon) =< q,\epsilon >$$

$$\widehat{\delta}(q,xa) = \{< p,w_1,w_2 > |\text{有}p'\in K\text{使} < p',w_1 >\in \widehat{\delta}(q,x)\text{且} < p,w_2 >\in \delta(p',a)\}$$

则$\widehat{\delta}$确实是一个从$K\times\Sigma^*$到$2^{K\times\Delta^*}$的函数,而且对任意的$q\in K, a\in\Sigma$及$x,y\in\Sigma^*$有:

$$\widehat{\delta}(q,a) = \delta(q,a)$$

$$\widehat{\delta}(q,xy) = \{< p,w_1w_2 > |\text{有}p'\in K,\text{使} < p',w_1 >\in \widehat{\delta}(q,x)\text{且} < p,w_2 >\in \widehat{\delta}(p',y)\}$$

即$\widehat{\delta}$确为δ的延拓,通常仍用δ表示$\widehat{\delta}$.

对例1中的$gsm\ M$,显然有:

$$\widehat{\delta}(q_0, 0^n 1^m) = \widehat{\delta}(q_0, 0^n) = a^{2(n-1)}b \qquad (m \geq 0, n \geq 1)$$

$$\widehat{\delta}(q_0, 1^m 0^n 1^k) = a^{m+2(n-1)}b \qquad (n \geq 1, m, k \geq 0)$$

$$\widehat{\delta}(q_0, 0^m 1^n) = \emptyset \qquad (m \geq 1, n \geq 0)$$

引 理5.1 设$M =< K, \Sigma, \Delta, \delta, q_0, F >$为$gsm$, 若对任意的$x \in \Sigma^*$及$w \in \Delta^*$,令:

$$M(x) = \{u \in \Delta^* | \text{有} q \in F, \text{使} < q, u > \in \delta(q_0, x)\}$$

$$M^{-1}(x) = \{y \in \Sigma^* | w \in M(y)\}$$

则M为从Σ^*到2^{Δ^*}的函数,M^{-1}为从Δ^*到2^{Σ^*}的函数.

并称M为gsm映射,M^{-1}为逆gsm映射.

引理5.2 设$M =< K, \Sigma, \Delta, \delta, q_0, F >$为$gsm$, 若令

$$R_M = \{[q_0, a_1, w_1, q_1][q_1, a_2, w_2, q_2] \cdots [q_{n-1}, a_n, w_n, q_n] | q_n \in F \text{且} < q_{i+1}, w_{i+1} > \in \delta(q_i, a_{i+1}), 0 \le i < n\}$$

则R_M为正规集.

证明： 令$\Sigma_M = \{[q, a, w, p] | < p, w > \in \delta(q, a)\}$

因为$\delta : K \times \Sigma \to 2^{K \times \Delta^*}$为有穷函数,而且$K$和$\Sigma$又都是有穷集,所以$\Sigma_M$也是有穷集.

1) 若$\Sigma_M = \emptyset$,则$R_M = \emptyset$,故R_M为正规集.

2) 若$\Sigma_M \neq \emptyset$,则定义部分函数$\delta' : K \times \Sigma_M \to K$如下:

当$[q, a, w, p] \in \Sigma_M$时,有

$\delta'(q, [q, a, w, p]) = p$

则$M' =< K, \Sigma_M, \delta', q_0, F >$为一个$\Sigma_M$ $_{fa}$且$T(M') = R_M$. 所以这时R_M也是正规集.

定理5.4.1 rl类和cfl类关于gsm映射封闭.

证明：设L为rl或cfl,且$gsm\ M =< K, \Sigma, \Delta, \delta, q_0, F >$.

令$L_1 = L \cap \Sigma^*$,则L和L_1同时为rl或者cfl.定义置换$f : \Sigma \longrightarrow 2^{\Sigma_M^*}$如下：

$f(a) = \{[q, a, w, p] | [q, a, w, p] \in \Sigma_M\}, a \in \Sigma$

再定义同态$h : \Sigma_M \longrightarrow \Delta^*$如下：

当$[q, a, w, p] \in \Sigma_M$时,有

$h([q, a, w, p]) = w$

则有$M(L) = M(L_1) = h(f(L_1) \cap R_M)$

由于rl和cfl类关于置换,同态及与正规集的交都封闭,所以L与$M(L)$同为rl或同为cfl.

定理5.4.2 rl类和cfl类关于逆gsm映射封闭.

证明： 设L为rl或cfl,且$grm\ M =< K, \Sigma, \Delta, \delta, q_0, F >$.

令$L_1 = L \cap \Delta^*$,则L_1与L同为rl,或同为cfl.

定义置换$f : \Delta \to (\Sigma \cup \Delta)^*$如下：

$f(b) = \Sigma^* b$

则每个$f(b)(b \in \Delta)$都是rl,当然也是cfl.再令

$$L_2 = \begin{cases} f(L_1) \cup \Sigma^* & \text{若}\epsilon \in L \\ f(L_1) & \text{否则} \end{cases}$$

显然,L_2与L_1同为rl,或同为cfl,而且$\epsilon \in L$当且仅当$\epsilon \in L_2$.这时,若分别定义同态$h_1 : \Sigma_M \to (\Sigma \cup \Delta)^*$或$h_2 : \Sigma \cup \Delta \to (\Sigma \cup \Delta)^*$如下：

$h_1([q, a, w, p]) = aw([q, a, w, p] \in \Sigma_M)$

$$h_2(u) = \begin{cases} u & u \in \Sigma \\ \epsilon & u \in \Delta \end{cases}$$

那么,当注意到

$h_1(R_M) = \{a_1w_1\ldots a_mw_m | a_1,\ldots,a_m \in \Sigma, w_1,\ldots,w_m \in \Delta^*\}$; 有$q_1,\ldots,q_{m-1} \in K$及$q_m \in F$, 使 $< q_{i+1}, w_{i+1} > \in \delta(q_i, a_{i+1}), 0 \le i < m\}$

时,显然有

$$M^{-1}(L) = M^{-1}(L_1) = h_2(L_2 \cap h_1(R_M))$$

但因rl类和cfl类关于置换,同态及正规集之交都封闭,而且R_M为正规集, 所以$M^{-1}(L)$与L同为rl,或同为cfl.

5.5. 程序设计语言不是cfl

研究cfl的许多诱因都来自于程序设计语言,所以自然要问,现有的程序设计语言是否为cfl?要了解这一问题,必须知道"什么是程序设计语言".

以$ALGOL$为例,它是精心设计的最早的程序设计语言之一,有效的$ALGOL$程序为满足以下条件的符号串的集合:

满足$ALGOL$的语法条件;

满足$ALGOL$的语义条件;

访问主页
标题页
◀◀ ▶▶
◀ ▶
第 60 页 共 88 页
返回
全屏显示
关闭
退出

定理5.5.1 $ALGOL_{60}$不是cfl.

证明：设L为有效$ALGOL_{60}$程序集合.若L为cfl,当令

$$\bar{L} = L \cap (\{begin\}\{real\}\{x\}^*\{;\}\{x\}^*\{:=\}\{x\}^*\{end\})$$

时,由于$\bar{L}$是一个cfl与一个rl之交,所以$\bar{L}$为cfl.但是, 另一方面,又有:

$$\bar{L} = \{begin\ \ real\ \ x^i;x^j := x^k\ \ end|i = j = k\text{且}i \geq 1\}$$

其中条件$i = j = k$是由语义规则"在有效的$ALGOL_{60}$程序中出现的每个变量必须要预先说明"推出来的.其次,$i \geq 1$是由于有效的$ALGOL_{60}$程序中标识符的长度是不受限制的.现在取$gsm\ M$如下图所示:

则有$M(L) = \{a^n b^n c^n | n \geq 1\}$,它不是$cfl$,这与定理5.4.1矛盾，所以$L$不是$cfl$.

访问主页
标题页
第 61 页 共 88 页
返回
全屏显示
关闭
退出

定理5.5.2 $FORTRAN$不是cfl.

证明： 设L为有效的$ASA\ FORTRAN$程序集合.由于它的相继输入或输出卡片数是不受限制的,当令

$$R = \{WRITE(6,1)$$
$$1\ \ FORMAT(1\}\{0\}^*\{H\}\{x\}^*\{)END\}$$

$$R = \{WRITE(6,1)$$
$$1\ \ FORMAT\ \ (10^n Hx^{10^n})END|n \geq 0\}$$

所以当定义同态ϕ如下：

$$\phi(x) = x$$

$$\phi(a) = \epsilon(a \neq x)$$

时,则$\phi(L \cap R)$应为cfl.但另一方面,$\phi(L \cap R) = \{x^{10^n}|n \geq 0\}$不是$cfl$,得到矛盾.所以$L$不是$cfl$.

证明： 设L为有效的$PL/1$程序集合.假定L为cfl.令

$$\bar{L} = L \cap (\{MAH : PROCEDURE\ OPTIONS(MAIN) : DECLARE\ \ A(1 : 4\}\{,1 : 4\}^*\{); A(2)\{,2\}^*\{) = A(3\}\{,3\}^*\{; END\ MAH; \})$$

因为括号内的量是正规集,所以$\bar{L}$为cfl.因此

$$\bar{L} = \{MAH : PROCEDURE\ OPTIONS(MAIN); DECLARE\ A(1 : 4\{,1 : 4\}^i); A(2\{,2\}^i) = A(3\{,3\}^i); END\ MAH; | i \geq 0\}$$

这是因为：

i) 在$PL/1$中数组的维数是不受限制的;

ii) 每提到数组时,都必有相同的下标数;

若定义同态ϕ如下：

$\phi(4) = a$

$\phi(2) = b$

$\phi(3) = c$

$\phi(x) = \epsilon$(若$x \notin \{2, 3, 4\}$)

5.6. 单字母表上的上下文无关语言

定理5.6.1 设$L \subseteq \{a\}^*$,则L为cfl,当且仅当L为rl.

证明: 充分性是显然的,下面只证必要性.

设$L \subseteq \{a\}^*$为上下文无关语言.根据$uvwxy$字定理,必有正整数n_0,当$a^n \in L$且有$n \geq n_0$时,有:

$a^n = uvwxy \quad vx \neq \epsilon$且$|vwx| \leq n_0$

使$uv^iwx^iy \in L(i = 0, 1, 2, \ldots)$

当记$n_1 = |uwy|$及$n_2 = |vx|$时,由上即知有

$a^n = a^{n_1+n_2}(1 \leq n_2 \leq n_0)$

使$a^{n_1+n_2} \in L(i = 0, 1, 2, \ldots)$

1) 若p为$1, 2, \ldots, n_0$的最小公倍数且$w \in L \cap a^{n_0} \cdot a^*$,则$w(a^p)^* \subseteq L$

实际上,若$\bar{w} \in w(a^p)^*$,则必有$i \geq 0$,使$\bar{w} = w(a^p)^i = wa^{ip}$

再由前面结果,就得到$\bar{w} = a^{n_1+n_2+ip}$

但p为$1, 2, \ldots, n_0$的最小公倍数,且$1 \leq n_2 \leq n_0$.所以$n_2|p$,即有正整数q,使$p = n_2q$.因此

$\bar{w} = a^{n_1+(1+iq)n_2}$且$1 + iq \geq 1$

所以$\bar{w} \in L$.这表明$w(a^p)^* \subseteq L$.

2) 令$A_i = a^{n_0+i} \cdot (a^p)^* \cap L(i = 1, \ldots, p)$

当$A_i \neq \emptyset$时,用y_i表示A_i中长度最短的字.

3) $L = (L \cap (\cup_{i=0}^{n_0} a^i)) \cup (\cup_{i=1_{A_i \neq \emptyset}}^{p} y_i(a^p)^*)$

首先,因为

$L \cap (\cup_{i=0}^{n_0} a^i) \subseteq L$;

由1)和2)还知道,$y_i(a^p)^* \subseteq L(1 \leq i \leq p$且$A_i \neq \emptyset)$

所以$(L \cap (\cup_{i=0}^{n_0} a^i)) \cup (\cup_{i=1_{A_i \neq \emptyset}}^{p} y_i(a^p)^*) \subseteq L$

其次,任取$x \in L$,则只有以下两种可能:

其一,$|x| \le n_0$

这时显然有$x \in L \cap (\cup_{i=0}^{n_0} a^i)$.

其二,$|x| > n_0$

这时必有正整数k,使$|x| = n_0 + k$,设

$k = qp + t(1 \le t \le p)$

则$x = a^{n_0+i}(a^p)^q$,即$x \in a^{n_0+t}(a^p)^*$.所以$A_i \neq \emptyset$且y_i存在.因此,有$s \ge 0$且$s \neq q$使$y_i = a^{n_0+i}(a^p)^s$.这表明$x = y_i(a^p)^{q-s}$

即$x \in \cup_{i=1_{A_i \neq \emptyset}}^{p} y_i(a^p)^*$

总之,由$i)$和$ii)$知道,恒有

$x \in (L \cap (\cup_{i=0}^{n_0} a^i)) \cup (\cup_{i=1_{A_i \neq \emptyset}}^{p} y_i(a^p)^*)$

这表明还有: $L \subseteq (L \cap (\cup_{i=0}^{n_0} a^i)) \cup (\cup_{i=1_{A_i \neq \emptyset}}^{p} y_i(a^p)^*)$

4) 因为 $L \cap (\cup_{i=0}^{n_0} a^i))$ 为有穷集,所以它为正规语言,又因为每个 $y_i(a^p)^*$(其 $A_i \neq \emptyset$)都是正规语言,所以 $(\cup_{i=1_{A_i \neq \emptyset}}^{p} y_i(a^p)^*)$ 也是正规语言,从而可知,$L = (L \cap (\cup_{i=0}^{n_0} a^i)) \cup (\cup_{i=1_{A_i \neq \emptyset}}^{p} y_i(a^p)^*)$ 为正规语言.

5.7. $Baker$定理

定义5.7.1 设$G=<V_N,V_T,P,S>$为0型文法,且$V=V_N\cup V_T$.如果P中每个产生式都有如下形式:

$$x_0A_1x_1\cdots x_{n-1}A_nx_n\rightarrow x_0\gamma_1x_1\cdots x_{n-1}\gamma_nx_n$$

其中$x_0,x_1,\cdots,x_n\in V_T^*$, $A_1,\cdots,A_n\in V_N$且$\gamma_1,\cdots,\gamma_n\in V^*$. 就称$G$为非确定的重写文法,简称$NR-$文法.

引理5.3 设$G = < V_N, V_T, P, S >$为NR−文法,$w \in V_T^*$且$|w| > max\{L_G, B_G, R_G\}$.则有$\alpha_0, \alpha_1, \cdots, \alpha_{n+1} \in V^*$,使:

$$\alpha_0 w \alpha_1 w \cdots \alpha_n w \alpha_{n+1} \overset{m}{\Rightarrow} u$$

当且仅当有$u_0, u_1, \cdots, u_{n+1} \in V^*$,使

$$u = u_0 w u_1 w \cdots u_n w u_{n+1}$$

及

$$\alpha_0 w \overset{m_0}{\Rightarrow} u_0 w, w\alpha_i w \overset{m_i}{\Rightarrow} w u_i w \text{且} w\alpha_{n+1} \overset{m_{n+1}}{\Rightarrow} w u_{n+1} \qquad (i = 1, 2, \cdots, n)$$

其中$m = m_0 + m_1 + \cdots + m_{n+1}$.

证明：充分性是显然的,下面只证必要性.用关于m的归纳法来证.

当$m=0$时,命题显然成立;

假定对任意的$k \geq 0$,当$m=k$时命题也成立;

当$m=k+1$时,由于G为$NR-$文法且$|w| > max\{L_G, R_G, B_G\}$.所以，第一步派生所用的产生式$\alpha \rightarrow \beta$中的$\alpha$必为$\alpha_0 w, w\alpha_{n+1}$或$w\alpha_i w(1 \leq i \leq n)$的一个子字.因此$\alpha_0 w \alpha_1 w \ldots \alpha_n w \alpha_{n+1} \overset{k+1}{\Rightarrow} u$必有如下形式:

$$\alpha_0 w a_1 w \ldots \alpha_n w \alpha_{n+1} \Rightarrow \alpha_0 w \alpha_1 w \ldots \alpha_{i-1} w \beta' w \alpha_{i+1} \ldots \alpha_n w \alpha_{n+1} \overset{k}{\Rightarrow} u$$

其中$0 \leq i \leq n+1, \beta' \in V^*$且$\beta'$由$\alpha$的形状所决定,可有$\beta' \neq \beta$.由归纳假设可知,必有$u_0, u_1, \ldots, u_{n+1} \in V^*$使

$$u = u_0 w u_1 w \ldots u_n w u_{n+1}$$

且$\alpha_0 w \overset{k_0}{\Rightarrow} u_0 w, w\alpha_j w \overset{k_j}{\Rightarrow} w u_j w(1 \leq j \leq n$且$j \neq i)$

$w\beta' w \overset{k_i}{\Rightarrow} w u_i w, w\alpha_{n+1} \overset{k_{n+1}}{\Rightarrow} w u_{n+1}(k = k_0 + k_1 + \cdots + k_{n+1})$

因此有$a\alpha_i w \Rightarrow w\beta' w \overset{k_i}{\Rightarrow} w u_i w$及$k+1 = k_0 + k_1 + \cdots + k_{n+1} + 1$

这表明,当$m=k+1$时,命题也成立.

定义5.7.2 设$G =< V_N, V_T, P, S >$为$NR-$文法且$V = V_N \cup V_T$.如果P中每个产生式都具有如下形式:

$$x_0 A_1 x_1 \cdots x_{n_1} A_n x_n \to y_0 B_1 y_1 \cdots y_{m_1} B_m y_m$$

其中$x_0, x_1, \cdots, x_n, y_0, y_1, \cdots, y_m \in V_T^*, A_1, \cdots, A_n, B_1, \cdots, B_m \in V_N$.此外,或者$n = 1$,或者有$j(0 \leq j \leq m)$使

$$|y_j| > |x_k| \qquad (k = 1, 2, \cdots, n-1)$$

就称G为终端受限文法,简记为tb文法.

对每个0型文法$G =< V_N, V_T, P, S >$,令

$$N(G) = \Sigma_{\alpha \to \beta \in P}(|\alpha| - 1)$$

显然,G为cfg当且仅当$N(G) = 0$.

引理5.4 设$G=<V_N, V_T, P, S>$为tb文法,$V=V_N \cup V_T$.如果$N(G)>0$,则有tb文法G_1, gsm映射$\tilde{S}$及正规集R使

$$L(G)=\tilde{S}(L(G_1)\cap R)\text{且}N(G_1)<N(G).$$

证明：任取两个不同的新符号$, ¢ $\notin V$,以下分两种情况讨论.

情况一：$L_G > B_G$或$R_G > B_G$

不失一般性,假定$R_G > B_G$且$R_G \geq L_G$.否则,只须考虑文法$G^R =< V_N, V_T, \{\alpha^R \to \beta^R | \alpha \to \beta \in P\}, S >$即可.

取$\gamma w \to \beta w \in P$使$w \in V_T^*$且$|w| = R_G$,并令

$G_1 =< V_N, V_T \cup \{\$\}, P_1, S >$

其中$P_1 = (P \setminus \{\gamma w \to \beta w\}) \cup \{\gamma \to \beta w\$\}$

因为G为tb文法,所以G_1也是tb文法且$N(G_1) < N(G)$.

取正规集$R = (V \cup \{\$w\})^*$.若$w = a_1 \ldots a_n$且$a_1, \ldots, a_n \in V_T$,则取$gsm$映射$\tilde{S}$如下图所示:

若$x \in L(G)$,则$S \stackrel{*}{\Rightarrow}_G x$且有$u_0, u_1, \ldots, u_k \in V^* \setminus V^*wV^*$,使

$x = u_0 w u_1 w \ldots u_{k-1} w u_k$

根据引理5.3,及归纳法可知,必有$i_1, i_2, \ldots, i_k \geq 0$,使

$S \stackrel{*}{\Rightarrow}_G u_0(w\$)^{i_1} w u_1 (w\$)^{i_2} w \ldots u_{k-1}(w\$)^{i_k} w u_k$

因此,$y = u_0(w\$)^{i_1} w u_1 (w\$)^{i_2} w \ldots u_{k-1}(w\$)^{i_k} w u_k \in L(G_1) \cap R$,且$x = \tilde{S}(y)$.

从而得到$x \in \tilde{S}(L(G_1) \cup R)$.

反之,若$y \in L(G_1) \cap R$,则有

$S \Rightarrow_{G_1} \alpha_1 \Rightarrow_{G_1} \alpha_2 \Rightarrow_{G_1} \cdots \Rightarrow_{G_1} \alpha_n$且$\alpha_n = y$

显然$\alpha_1, \alpha_2, \ldots, \alpha_n \in R$.所以当记$\alpha_0 = S$时,由$P_1$的定义即可得到

$\tilde{S}(\alpha_i) \stackrel{*}{\Rightarrow}_G \tilde{S}(\alpha_{i+1}) (0 \leq < n)$

所以$S \stackrel{*}{\Rightarrow}_G \tilde{S}(y)$,即$S(y) \in L(G)$.

总结以上结果,即得到$L(G) = \tilde{S}(L(G_1) \cap R)$.

情况二　$L_G \le B_G$且$R_G \le B_G$

取G的一个产生式

$x_0 A_1 x_1 \ldots x_{n-1} A_n x_n \to x_0 \gamma_1 x_1 \ldots x_{n-1} \gamma_n x_n$

其中

1) $x_0, x_1, \ldots, x_n \in V_T^*, A_1, A_2, \ldots, A_n \in V_N$且$\gamma_1, \ldots, \gamma_n \in V^*$.

有$j(1 \le j < n)$,使$B_G = |x_j|$.

若$n = 1$,则$B_G = 0$.因此$L_G = R_G = 0$,从而可得$N(G) = 0$.这与题设$N(G) > 0$矛盾,所以必有$n \ge 2$.又因为G是tb文法,所以这个产生式右端必有字$w \in V_T^*$,使$|w| > B_G$.但$B_G \ge L_G$且$B_G \ge R_G$,故而$|w| > |x_i|(1 \le i \le n)$.这样,必有某个$\gamma_i(1 \le i \le n)$使$\gamma_i$或它的一部分包含在$w$中,因此,这个产生式必有如下形式:

$\alpha A_k x_k A_{k+1} \rho \to \beta_1 w \beta_2$

其中$|\beta_1| < |x_0 \gamma_1 x_1 \ldots \gamma_{k-2} x_{k-1} \gamma_{k-1}|$

或$|\beta_2| < |\gamma_k x_{k+1} \gamma_{k+1} \ldots x_{n-1} \gamma n x_n|$

不妨设为$|\beta_1| < |x_0\gamma_1 x_1 \ldots \gamma_{k-2} x_{k-1} \gamma_{k-1}|$.令

$G_1 =< V_N, V_T \cup \{\$, \not\subset\}, P_1, S >$

$P_1 = (P \setminus \{\alpha A_k x_k A_{k+1}\rho \rightarrow \beta_1 w \beta_2\}) \cup \{\alpha A_k \rightarrow \beta_1 w \$, x_k A_{k+1}\rho \rightarrow x_k \not\subset w\beta_2\}$

因为G为tb文法,所以G_1也是tb文法.而且还有$N(G_1) < N(G)$.

访问主页
标题页
第 76 页 共 88 页
返回
全屏显示
关闭
退出

取正规集$R=(V\cup\{\,\$x_k\not\subset w\})^*$.如果

$x_k=a_1\ldots a_p, w=b_1\ldots b_t$且$a_1,\ldots,a_p,b_1,\ldots,b_t\in V_T$

则取$gsm\ \tilde{S}$如图所示:

若$x\in L(G)$,则$S\Rightarrow_G x$,且有$u_0,u_1,\ldots,u_k\in V^*\setminus V^*wV^*$,使

$x=u_0wu_1w\ldots u_{k-1}wu_k$

根据引理5.3,由归纳法知,必有$i_1,i_2,\ldots,i_h\geq 0$,使

$S\overset{*}{\Rightarrow}_{G_1} u_0w(\$x_k\not\subset w)^{i_1}u_1\ldots u_{k-1}w(\$x_k\not\subset w)^{i_k}u_k$

因此,$y=u_0w(\$x_k\not\subset w)^{i_1}u_1\ldots u_{k-1}w(\$x_k\not\subset w)^{i_k}u_k\in L(G_1)\cap R$.所以由$x=S(y)$知道, $x\in\tilde{S}(L(G_1)\cap R)$.

反之,任取$y \in L(G_1) \cap R$,则有

$S \Rightarrow_{G_1} \alpha_1 \Rightarrow_{G_1} \alpha_2 \Rightarrow_{G_1} \cdots \Rightarrow_{G_1} \alpha_n$且$\alpha_n = y$

为了证明$S \overset{*}{\Rightarrow}_G \tilde{S}(y)$,只须先证明以下的命题:

如果$\theta, \beta \in R$且$\theta \overset{*}{\Rightarrow}_{G_1} \beta$,则有派生

$\theta \Rightarrow_{G_1} \theta_1 \Rightarrow_{G_1} \theta_2 \Rightarrow_{G_1} \cdots \Rightarrow_{G_1} \theta_m$且$\theta_m = \beta$

则产生式$\alpha A_k \rightarrow \beta_1 w\$$和$x_k A_{k+1}\rho \rightarrow x_k \not\subset w\beta_2$都相继被使用.

用在派生$\theta \overset{*}{\Rightarrow}_{G_1} \beta$中产生式$\alpha A_k \rightarrow \beta_1 w\$$和$x_k A_{k+1}\rho \rightarrow x_k \not\subset w\beta_2$不相继被使用的次数$m$之归纳法来证.

当$m = 0$时,命题显然为真;

假定对任意的自然数k,当$m = k$时命题为真.

当$m = k + 1$时,第一次不相继使用的产生式为$\alpha A_k \rightarrow \beta_1 w\$$或$x_k A_{k+1}\rho \rightarrow x_k \not\subset w\beta_2$

首先,若为$\alpha A_k \rightarrow \beta_1 w\$$.这时$\theta \overset{*}{\Rightarrow}_{G_1} \beta$必有如下形式:

$\theta \overset{*}{\Rightarrow}_{G_1} \theta_1 \alpha A_k x_k A_{k+1}\rho\theta_2 \Rightarrow_{G_1} \theta_1\beta_1 w\$ x_k A_{k+1}\rho\theta_2 \overset{*}{\Rightarrow}_{G_1} \beta$

因为\$不在G_1中任何产生式的左端出现.所以由引理5.3知道,必有$u_1, u_2 \in V^*$使$\beta = u_1 w\$ x_k u_2, \theta_1\beta_1 w \overset{*}{\Rightarrow}_{G_1} u_1 w$且$x_k A_{k+1}\rho\theta_2 \overset{*}{\Rightarrow}_{G_1} x_k u_2$.但因$u_1 w\$ x_k u_2 \in \boldsymbol{R}$,所以必有$u_2' \in V^*$,使$u_2 =\not\subset w u_2'$.

从而知道,$x_k A_{k+1}\rho\theta_2 \stackrel{*}{\Rightarrow}_{G_1} x_k \not\subset wu_2'$必有如下形式:

$$x_k A_{k+1}\rho\theta_2 \stackrel{*}{\Rightarrow}_{G_1} x_k A_{k+1}\rho\tilde{u}_2 \Rightarrow_{G_1} x_k \not\subset w\beta_2\tilde{u}_2 \stackrel{*}{\Rightarrow}_{G_1} x_k \not\subset wu_2'$$

从而得到

$$\theta \stackrel{*}{\Rightarrow}_{G_1} \theta_1\alpha A_k x_k A_{k+1}\rho\theta_2 \stackrel{*}{\Rightarrow}_{G_1} \theta_1\alpha A_k x_k A_{k+1}\rho\tilde{u}_2 \Rightarrow_{G_1} \theta_1\beta_1 w\$ x_k A_{k+1}\rho\tilde{u}_2 \Rightarrow_{G_1} \theta_1\beta_1 w\$ x_k \not\subset w\beta_2\tilde{u}_2 \stackrel{*}{\Rightarrow}_{G_1} \beta$$

从而由归纳假设即知,当$m = k + 1$时命题也真.

其次,若为$x_k A_{k+1}\rho \to x_k \not\subset w\beta_2$.可以通过和前面$\alpha A_k \to \beta_1 w\$$时类似的证明得知,此时命题也真.于是就证明了命题.

根据上面刚证明的命题,不难由归纳法得到,如果$\theta, \beta \in R$且$\theta \stackrel{*}{\Rightarrow}_{G_1} \beta$,则$\tilde{S}(\theta) \stackrel{*}{\Rightarrow} \tilde{S}(\beta)$.从而得到

$S \stackrel{*}{\Rightarrow} \tilde{S}(y)$ 即$\tilde{S} \in L(G)$.

总结以上所述,在$B_G \geq L_G$且$B_G \geq R_G$时,也有

$L(G) = \tilde{S}(L(G_1) \cap R)$.

定理5.7.1($Baker$定理) 设$G =< V_N, V_T, P, S >$为tb文法,则$L(G)$为上下文无关语言.

证明: 用关于$n = N(G)$的第二归纳法来证明.

当$n = 0$时,G为cfg,所以$L(G)$为cfl.

当$n = k$时,根据引理5.4,有tb文法G_1, gsm映射$\tilde{S}$及正规集R使$L(G) = \tilde{S}(L(G_1) \cap R)$且$N(G_1) < N(G)$.根据归纳假设$L(G_1)$为$cfl$.从而知道$L(G) = \tilde{S}(L(G_1) \cap R)$也必为$cfl$.

定理5.7.2 若0型文法$G=<V_N, V_T, P, S>$的产生式都具有如下形式

$$\alpha \to \beta \quad \alpha \in V_N^+ \text{且} \beta \in V^* V_T V^*$$

其中$V = V_N \cup V_T$,则$L(G)$为cfl.

定理5.7.3 若0型文法$G = < V_N, V_T, P, S >$的产生式都具有如下形式

$$xAy \to x\beta y \quad x, y \in V_T^*, A \in V_N \text{且} \beta \in (V_N \cup V_T)^*$$

则$L(G)$为cfl.

定理5.7.4 若0型文法$G=<V_N,V_T,P,S>$的产生式都具有如下形式

$$xA\beta \to x\gamma\beta \quad x\in V_T^*, A\in V_N\text{且}\gamma,\beta\in V^*$$

则当$|x|\geq|\beta|$时,$L(G)$必为cfl.

定理5.7.5 若0型文法$G=< V_N, V_T, P, S >$中的产生式都具有如下形式

i) $xA\beta \to x\gamma\beta$且$|x| \geq |\beta|$

或

ii) $\alpha Ay \to \alpha\gamma y$且$|y| \geq |\alpha|$

则$L(G)$必为cfl.

定义5.7.3 设$G=<V_N,V_T,P,S>$为0型文法,令;

$$Left(G)=\{w\in V_T^*|S\overset{*}{\Rightarrow}_L w\}$$

$$Right(G)=\{w\in V_T^*|S\overset{*}{\Rightarrow}_R w\}$$

$$Two-way(G)=\{w\in V_T^*|S\overset{*}{\Rightarrow}_{L,R} w\}$$

定理5.7.6 设$G=<V_N,V_T,P,S>$为$NR-$文法,则$Left(G),Right(G),Two-way(G)$都是$cfl$.

设$V=V_N\cup V_T$,任取四个不同的新符号d,L,R和S_1,令

$G'=<V_N\cup\{L,R,S_1\},V_T\cup\{d\},P_1,S_1>$

其中P_1由以下两组产生式组成:

1) 当记$q=max\{|\alpha|\alpha\rightarrow\beta\in P\}$时,$P_1$包含以下的产生式:

$S_1\rightarrow d^qLSRd^q$

$L\rightarrow\epsilon$

$R\rightarrow\epsilon$

2) 若$u\alpha v\rightarrow w\beta x\in P,u,v,w,x,y\in V_T^*,|y|=q$且$\alpha,\beta\in\{\epsilon\}\cup V_NV^*$时,则$P_1$包含以下的产生式:

$yuLav\rightarrow ywL\beta x$

$u\alpha Rvy\rightarrow w\beta Rxy$

$yuL\alpha Rv\rightarrow ywL\beta Rxy$

显然,L和R分别指出派生中的最左和最右的非终极符号,而且在派生中,每步都仅能对含最左或最右的非终极符号之部分进行重写,因此有

$L(G')=Two-way(G')$

但G'为tb文法,所以$L(G')$为cfl.若定义同态h如下:

$h(d)=\epsilon$

定理5.7.7 设$G=<V_N,V_T,P,S>$为$NR-$文法,$V=V_N\cup V_T$且$<$为V上的一个半序,若G中每个产生式都有如下形式:

$$A_1\cdots A_n\to B_1\cdots B_m$$

其中,$A_1\cdots A_n\in V_N$,且有$k(1\le k\le m)$使得对每个$i(1\le i\le n)$皆有$A_i<B_k$,则$L(G)$为cfl.

证明:取$p=max\{|\beta||\alpha\to\beta\in P\}$及三个不同的新符号$d,E,S_1$.把$V$上的半序$<$延拓为$V\cup\{d,E,S_1\}$上的一个全序,使$S_1$为最小元,$d$为最大元,且$E$为次大元. 若定义同态$h:V\cup\{d\}\to V^*$如下:

$$h(C)=\begin{cases}\epsilon & C=d\\ A_{p^i} & C\text{为全序}<\text{之下第}i\text{个元素}\end{cases}$$

就可取$G_1=<V_N\cup\{E,S_1\},V_T\cup\{d\},P_1,S_1>$

其中$P_1=\{h(\alpha)Y\to h(\beta)Y|\alpha\to\beta\in P\text{且}Y\in V_N\}\cup\{S_1\to SdE,E\to d\}$

则G_1为tb文法,且$h(L(G_1))=L(G)$.所以$L(G)$为cfl.