第十一讲

上下文无关语言的性质

上下文无关语言的性质

- CFL的必要条件
- CFL的Pumping引理
- CFL的闭运算性质
- CFL的同态性质
- CFL的交运算

上下文无关语言的性质

- CFL的必要条件
- CFL的Pumping引理
- CFL的闭运算性质
- CFL的同态性质
- CFL的交运算

观察

- Chomsky 范式(CNF)具有很好的文法结构；
- 任何CFG都能容易地转化为CNF；
- CNF可以有效地用于文法解析和定理证明。

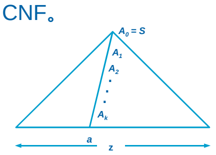
CFL的必要条件

设 L 为不含 ε 的无限CFL，其CFG:
$G = (V, T, S, P)$ 为 CNF。

记 $|V|=m$，$n=2^m$.

对 $z \in L$, 考虑 z 的分析树。
若最大路径的长度 m, 则
$|z| \leq 2^{m-1} < n$.

见右图。



CFL的必要条件

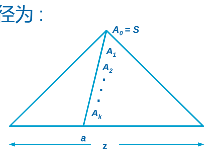
若 $|z| \geq n = 2^m$, 则 分析树的高度至少为$m+1$。
设 z 从S开始的分析树最大路径为：

$A_0A_1A_2 \ldots A_ka$

则有, $k \geq m$
然而, $|V|=m$, 因此,

$A_{k-m}, A_{k-m+1}, \ldots, A_{k-1}, A_k$

中必然有相同的变量. 设 $A_i = A_j$, 其中 $k\text{-}m \leq i < j \leq k$



CFL的必要条件

z 的分析树见下图. z 可以表示为 $z=uvwxy$. 其中：

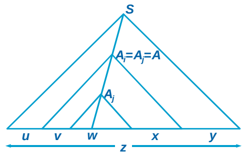
w 是根结点为A_j子树的产物。

vwx 是根结点A_i的产物,v 和 x 分别位于串w 的左右。

由于没有单一产生式, 因而 $vx \neq \varepsilon$

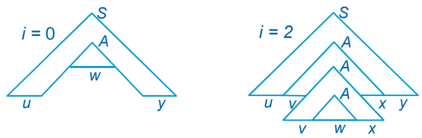
又根结点为A_i的子树最高为 $m+1$

则 $|vwx| \leq 2^m = n$



CFL的必要条件

对 v 和 x 得到"pumping"性质。如下图是 i=0,2 的情形。



"pumping" 的特性: 对任意 $i \geq 0$, $uv^iwx^iy \in L$

这为无限CFL的必要条件

上下文无关语言的性质

- CFL的必要条件
- CFL的Pumping引理
- CFL的闭运算性质
- CFL的同态性质
- CFL的交运算

CFL泵引理

定理：设 L 是 CFL. 则存在常数 n, 对任意 $z \in L$, $|z| \geq n$, z 可表示为 $z=uvwxy$, 且满足：

1. $vx \neq \varepsilon$
2. $|vwx| \leq n$
3. 对任意 $k \geq 0$, $uv^kwx^ky \in L$

证明：L 是 $\varnothing$ 或 $\{\varepsilon\}$结论显然成立.

否则，设CFG $G = (V, T, S, P)$ 为 CNF，且语言为 $L-\{\varepsilon\}$，根据前面的讨论，取 $n = 2^{|V|}$ 即可.

CFL泵引理

- 泵引理的条件可形式化的表示为:

$$\exists n \forall z \exists u \exists v \exists w \exists x \exists y \forall k (z \in L \wedge |z| \geq n \rightarrow z=uvwxy \wedge vx \neq \varepsilon \wedge |vwx| \leq n \wedge (k \geq 0 \rightarrow uv^kwx^ky \in L))$$

- 泵引理的否命题为:

$$\forall n \exists z \forall u \forall v \forall w \forall x \forall y \exists\ k (z \in L \wedge |z| \geq n \wedge (z=uvwxy \wedge vx \neq \varepsilon \wedge |vwx| \leq n \rightarrow k \geq 0 \wedge uv^kwx^ky \notin L))$$

用于证明 L 不是CFL

CFL泵引理的应用

证明不是CFL的步骤:

1. 选取任意正整数 n
2. 找一个字符串 $z \in L$, 使得:
   (i) $|z| \geq n$
   (ii) 对满足如下条件的任意 u,v,w,x,y
   $z=uvwxy$, $vx \neq \varepsilon$, $|vwx| \leq n$ 注意只能保证vx一个不为空
   (iii) 选择 $k \geq 0$, 使得
   $uv^kwx^ky \notin L$

CFL泵引理的应用

例：
证明语言
$L_{012} = \{ 0^k 1^k 2^k \mid k \geq 1 \}$
不是 CFL

CFL泵引理的应用

证明: $L_{012} = \{ 0^k 1^k 2^k \mid k \geq 1 \}$
对任意正整数 n
取 $z = 0^n 1^n 2^n \in L_{012}$
对满足 $z=uvwxy$, $vx \neq \varepsilon$, $|vwx| \leq n$ 的任意 u, v, w, x, y
显然 v, w, x 不可能取得全部的 0, 1, 2
若 k=0, 则, $uv^kwx^ky = uwy \notin L_{012}$
根据泵引理，语言L_{012} 不是CFL。

CFL泵引理的应用

例：证明语言
$L = \{ ww \mid w \in \{0,1\}^* \}$
不是 CFL.

CFL泵引理的应用

证明: $L = \{ ww \mid w \in \{0,1\}^* \}$
对任意正整数 n
取 $z = 0^n 1^n 0^n 1^n \in L$
对满足 $z=uvwxy$, $vx \neq \varepsilon$, $|vwx| \leq n$的任意 u, v, w, x, y
若 k=0, 则,
$uv^kwx^ky = uwy \notin L$
根据泵引理，语言L不是CFL。

上下文无关语言的性质

- CFL的必要条件
- CFL的Pumping引理
- CFL的闭运算性质
- CFL的同态性质
- CFL的交运算

CFL的闭运算性质

CFL的基本运算：
- 并
- 闭包 ((*), (+))
- 连接
- 反转
- 交
- 补

CFL的闭运算性质

CFL的并：

L 是CFL
M 是CFL
➡ ? $L \cup M$ 是CFL

定理:
若 L 和 M 都是 CFL, 则 $L \cup M$ 也是CFL。

CFL的闭运算性质

例：

语言	CFG
$L_1 = \{a^n b^n\}$	$S_1 \to aS_1b \mid \varepsilon$
$L_2 = \{ww^R\}$	$S_2 \to aS_2a \mid bS_2b \mid \varepsilon$

并

$L = \{a^n b^n\} \cup \{ww^R\}$ $\quad S \to S_1 \mid S_2$

CFL的闭运算性质

一般地：

对于CFL	L_1, L_2
对应CFG	G_1, G_2
开始变量	S_1, S_2
CFL的并	$L_1 \cup L_2$
新的开始变量	S
另加产生式	$S \to S_1 \mid S_2$

CFL的闭运算性质

CFL的闭包：

L 是CFL ➡ ? L^* 是CFL, L^+ 是CFL

定理:
若 L 是 CFL, 则 L^* 和 L^+ 都是CFL。

CFL的闭运算性质

例：

语言	CFG
$L_1 = \{a^n b^n\}$	$S \to aSb \mid \varepsilon$

星闭包

$L = \{a^n b^n\}^*$ $\quad S_1 \to SS_1 \mid \varepsilon$

CFL的闭运算性质

一般地：

对于CFL	L
对应CFG	G
开始变量	S
CFL的星闭包	L^*
新的开始变量	S_1
另加产生式	$S_1 \to SS_1 \mid \varepsilon$

CFL 的连接

CFL的连接：

L 为CFL
M 为CFL

? LM 是CFL

定理:
若 L 和 M 都是 CFL, 则LM 也是CFL 。

CFL 的连接

例：

语言	文法
$L_1=\{a^nb^n\}$	$S_1 \to aS_1b \mid \varepsilon$
$L_2=\{ww^R\}$	$S_2 \to aS_2a \mid bS_2b \mid \varepsilon$

连接

$L=\{a^nb^n\}\{ww^R\}$ $\qquad S \to S_1S_2$

CFL 的连接

一般地:

对CFL	L_1, L_2
文法CFG	G_1, G_2
开始变量	S_1, S_2
语言的连接	L_1L_2
新的开始变量	S
增加产生式	$S \to S_1S_2$

CFL的反转

若 $w=a_1a_2\ldots a_n$, 则w 的反转为:

$w^R = a_na_{n-1}\ldots a_1$

语言L 的反转为:

$L^R = \{ w^R \mid w \in L \}$.

L 为 CFL ? L^R为 CFL

定理: 若 L 是 CFL, 则L^R 也是CFL.

CFL的反转

证明方法:

设 $L=L(G)$, 对应的CFG $G=(V, T, S, P)$. 构造

$G^R = (V, T, S, P^R)$

其中

$P^R = \{A \to \alpha^R \mid \text{“}A \to \alpha\text{”} \in P\}$

可以证明, $L(G^R) = L^R$, 即对任意 w,

$S \overset{*}{\underset{G}{\Rightarrow}} w$ iff $S \overset{*}{\underset{G_R}{\Rightarrow}} w^R$

(作为练习)

上下文无关语言的性质

- CFL的必要条件
- CFL的Pumping引理
- CFL的闭运算性质
- CFL的同态性质
- CFL的交运算

CFL的替换

- 定义: 设 Σ 为字母表, $\mathscr{L}$ 为一语言的集合。
  映射: $s : \Sigma \to \mathscr{L}$ 称为Σ上的一个替换, 即
  对任意 $a \in \Sigma$, $s(a) \in \mathscr{L}$ 为一语言。
- 替换扩展： $s : \Sigma^* \to \mathscr{L}$
  若 $w = a_1 a_2 \ldots a_n \in \Sigma^*$, 定义：
  $s(w) = s(a_1 a_2 \ldots a_n) = s(a_1) s(a_2) \ldots s(a_n)$;
- 设 L 为 Σ上的语言, 定义
  $s(L) = \cup_{w \in L}\, s(w)$

CFL的替换

例: 设 $\Sigma = \{0,1\}$ 上替换 s :

$s(0) = \{a^n b^n \mid n \geq 1\}$, $s(1) = \{aa, bb\}$

若 w =01, 则:

$s(w) = s(0)s(1)$

$= \{a^n b^n\, aa \mid n \geq 1\} \cup \{a^n b^{n+2} \mid n \geq 1\}$

若 $L = L(0^*)$, 则:

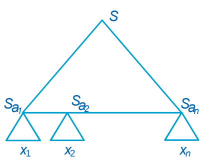
$s(L) = (s(0))^* = \{a^n b^n \mid n \geq 1\}^*$

$= \{a^{n_1} b^{n_1}\, a^{n_2} b^{n_2} \ldots a^{n_k} b^{n_k} \mid k \geq 0,\ n_i \geq 1,\ 1 \leq i \leq k\}$

CFL的替换

定理:

设 s 为 Σ上替换, 满足
$\forall a \in \Sigma$, $s(a)$ 是CFL。
若 L 是Σ 上的CFL, 则
$s(L)$ 是一个CFL。



证明方法:

参见右图所示的分析树。
$w = a_1 a_2 \ldots a_n$ 对应的分析树中每个叶结点 a ，
可替换为语言$s(a)$ 中任何串的分析树。

CFL的同态

设 $h : \Sigma \to T^*$, $w = a_1 a_2 \ldots a_n \in \Sigma^*$, 记

$h(w) = h(a_1)\, h(a_2) \ldots h(a_n)$

则 h 是一个同态映射。

对语言$L \subseteq \Sigma^*$,

$h(L) = \{ h(w) \mid w \in L \}$

则 h 是L 的一个同态映射.

定理: 若 L 为CFL，$h : \Sigma \to T^*$ 为同态映射, 则 $h(L)$ 也是CFL。

CFL的逆同态

设同态映射 $h : \Sigma \to T^*$, 对语言 $L \subseteq T^*$, 则 L 的逆同态 (*inverse homomorphism*) 为:

$h^{-1}(L) = \{ w \mid w \in \Sigma^*, h(w) \in L \}$

定理: 设 $L \subseteq T^*$ 是 CFL, $h : \Sigma \to T^*$ 为同态映射。
则 $h^{-1}(L)$ 也是 CFL。

证明方法：

设 $L = L(P)$, 其中 *PDA*

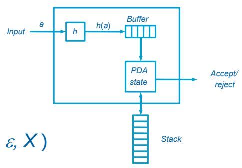
$P = (Q, T, \Gamma, \delta, q_0, Z_0, F)$

CFL的逆同态

构造 *PDA* $P' = (Q', \Sigma, \Gamma, \delta', (q_0,\varepsilon), Z_0, F \times \{\varepsilon\})$

其中 $Q' = Q \times \{ x \mid x$ 是$h(a)$ 的前缀, $a \in \Sigma \}$

对 $a \in \Sigma$,
$\delta'((q,\varepsilon), a, X)$
$= \{((q, h(a)), X)\}$
若$b \in T$ 或 $b = \varepsilon$,
$(p, \gamma) \in \delta(q, b, X)$,
则$((p,x), \gamma) \in \delta'((q,bx), \varepsilon, X)$



可以证明 $h^{-1}(L) = L(P')$

上下文无关语言的性质

- CFL的必要条件
- CFL的Pumping引理
- CFL的闭运算性质
- CFL的同态性质
- CFL的交运算

CFL的交运算

CFL的交运算：

L 是CFL，M 是CFL $\Rightarrow$? $L \cap M$ 是CFL

结论：

若 L 和 M 为CFL，但 $L \cap M$ 不一定是CFL。

CFL的交运算

$L_1 = \{a^n b^n c^m\}$

CFG：

$S \to AC$

$A \to aAb \mid \varepsilon$

$C \to cC \mid \varepsilon$

$L_2 = \{a^n b^m c^m\}$

CFG：

$S \to AB$

$A \to aA \mid \varepsilon$

$B \to bBc \mid \varepsilon$

交

$L_1 \cap L_2 = \{a^n b^n c^n\}$ 不是 CFG

CFL 的补和差

推论：

若 L 和 M 为CFL，但 $\overline{L}$ 和 $L\text{–}M$ 不一定是CFL 。

证明：

由于 $L \cap M = \overline{\overline{L} \cup \overline{M}}$ ，所以，CFL 的补运算不是封闭的。

由于 $\overline{L} = \Sigma^*\text{–}L$，所以，CFL 之间的差运算不是封闭的。

CFL 与正则语言的交

定理：

若 L 为CFL, R 为正则语言，则 $L \cap R$ 为CFL。

证明思路：

设 $R = L(A)$, 其中 DFA

$A = (Q_A, \Sigma, \delta_A, q_A, F_A)$

设 $L = L(P)$, 其中 PDA

$P = (Q_P, \Sigma, \Gamma, \delta_P, q_P, Z_0, F_P)$

构造 PDA：

$P' = (Q_P \times Q_A, \Sigma, \Gamma, \delta, (q_P, q_A), Z_0, F_P \times F_A)$

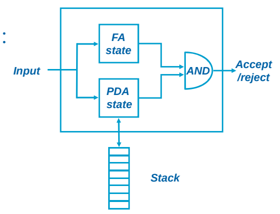
CFL 与正则语言的交

$\delta((q,p), a, X)$ 包含所有满足如下条件的 $((r,s),\gamma)$：

(1) $(r, \gamma) \in \delta_P(q, a, X)$

(2) $s = \delta_A^*(p, a)$

其中 $a \in \Sigma$ 或 $a = \varepsilon$



可证 $L \cap R = L(P')$

CFL 与正则语言的交

PDA M_1 CFL L_1

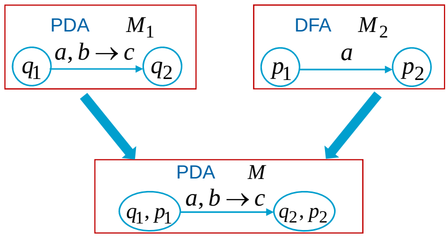
DFA M_2 正则语言 L_2

构造接受语言 $L_1 \cap L_2$ 的PDA M

M 并行模拟 M_1和 M_2

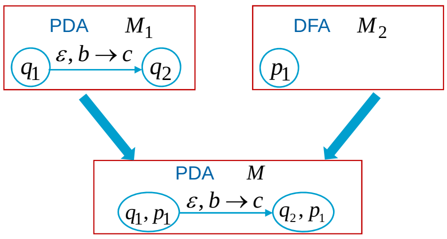
CFL 与正则语言的交

转移



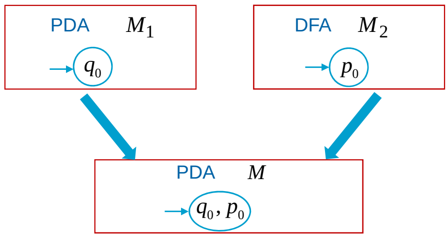
CFL 与正则语言的交

转移



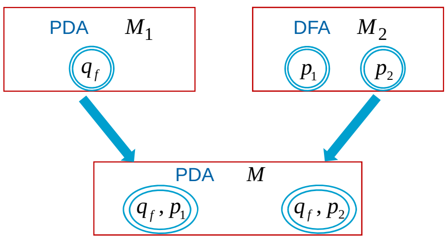
CFL 与正则语言的交

初态



CFL 与正则语言的交

终态

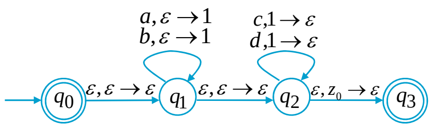


CFL 与正则语言的交

例:

CFL $L_1 = \{w_1 w_2 \mid |w_1| = |w_2|, w_1 \in \{a,b\}^*, w_2 \in \{c,d\}^*\}$

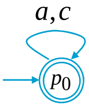
PDA M_1



CFL 与正则语言的交

正则语言 $L_2 = \{a, c\}^*$

DFA M_2

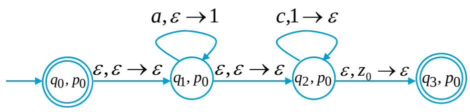


CFL 与正则语言的交

语言交的CFL:

$$L_1 \cap L_2 = \{a^n c^n \mid n \geq 0\}$$

PDA M



CFL性质的应用

$$\left.\begin{array}{l} L_1 \ \text{CFL} \\ L_2 \ \text{正则} \end{array}\right\} \Rightarrow L_1 \cap L_2 \ \text{CFL}$$

正则闭性质: CFL与正则语言的交为CFL

CFL性质的应用

证明:

$L = \{a^n b^n \mid n \neq 100, n \geq 0\}$ 是CFL

已知: $\{a^n b^n \mid n \geq 0\}$ 是CFL

显然: $L_1 = \{a^{100} b^{100}\}$ 是正则

$\Downarrow$

$\overline{L_1} = \{(a+b)^*\} - \{a^{100} b^{100}\}$ 也是正则的

CFL性质的应用

$\{a^n b^n\}$ $\qquad \overline{L_1} = \{(a+b)^*\} - \{a^{100} b^{100}\}$

CFL $\searrow$ $\qquad \swarrow$ 正则

$\{a^n b^n\} \cap \overline{L_1}$ CFL

$\Downarrow$

$\{a^n b^n\} \cap \overline{L_1} = \{a^n b^n \mid n \neq 100, n \geq 0\} = L$是CFL

CFL性质的应用

证明:

$$L = \{w \in \{a, b, c\}^* \mid n_a(w) = n_b(w) = n_c(w)\}$$

不是CFL

CFL性质的应用

若 $L = \{w \in \{a,b,c\}^* \mid n_a = n_b = n_c\}$ 是CFL

则 $L \cap \{a^*b^*c^*\} = \{a^n b^n c^n\}$

CFL 正则 CFL 不可能!!!

因此, *L* 不是CFL

课后练习

- 必做题:
  - P-280 *Ex.7.2.1*(*b*), (*c*), (*e*)
  - P-292 *!*Ex.7.3.1*(*b*)
  - P-292 *Ex.7.3.2*
  - P-293 *Ex.7.3.6*
- 思考题:
  - P-280 !*Ex.7.2.1*(*f*)
  - P-292 !!*Ex.7.3.3*(*b*), (*c*)

That's all for today.

Thank You