形式语言与自动机理论

上下文无关语言的性质

王春宇

计算机科学与技术学院
哈尔滨工业大学

上下文无关语言的性质

- 上下文无关语言的泵引理
  - 上下文无关语言的泵引理
  - 泵引理的应用
- 上下文无关语言的封闭性
- 上下文无关语言的判定性质
- 乔姆斯基文法体系

上下文无关语言的泵引理

定理 33

如果语言 L 是 CFL, 那么存在正整数 N, 对 $\forall z \in L$,
只要 $|z| \geq N$, 就可以将 z 分为五部分 $z = uvwxy$ 满足:

❶ $vx \neq \varepsilon \quad$ (或$|vx| > 0$);

❷ $|vwx| \leq N$;

❸ $\forall i \geq 0,\ uv^iwx^iy \in L$.

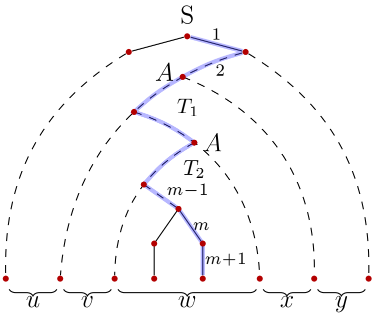
证明:

❶ 设 CNF 格式 CFG G 中变元数 $|V| = m$, 令 $N = 2^m$, 若有 $z \in L(G)$, 且 $|z| \geq N$.

❷ 则 z 的派生树内节点是二叉树, 最长路径长度至少 $m+1$, 节点至少 $m+2$ 个.

❸ 该路径由下至上 $m+1$ 个内节点中, 必有两个 T_2 和 T_1 标记了相同的变元 A.

❹ 若记 T_2 产物为 w, 且是 T_1 的子树, T_1 的产物可记为 vwx, 则有 $A \overset{*}{\Rightarrow} vAx$ 和 $A \overset{*}{\Rightarrow} w$.



❺ 那么 $\forall i \geq 0$, $A \overset{*}{\Rightarrow} v^i w x^i$. 不妨设 $z = uvwxy$, 则 $S \overset{*}{\Rightarrow} uAy \overset{*}{\Rightarrow} uv^i wx^i y$.

❻ T_1 路径长不超过 $m+1$, 那么 T_1 产物长不超过 2^m, 所以 $|vwx| \leq 2^m$.

❼ T_2 必在 T_1 的左/右儿子中, 所以 v 和 x 不可能同时为空, 即 $vx \neq \varepsilon$.

$\square$

泵引理的应用

例 1. 证明 $L = \{0^n1^n2^n \mid n \geq 1\}$ 不是上下文无关语言.

证明:

1. 假设 L 是 CFL, 那么存在整数 N, 对 $\forall z \in L\,(|z| \geq N)$ 满足泵引理.
2. 从 L 中取 $z = 0^N1^N2^N$, 则显然 $z \in L$ 且 $|z| = 3N \geq N$.
3. 由泵引理, z 可被分为 $z = uvwxy$, 且有 $|vwx| \leq N$ 和 $vx \neq \varepsilon$.
4. 那么 vwx 可能
   1. 只包含 0, 1 或 2, 那么 $uwy \notin L$;
   2. 只包含 0 和 1, 或只包含 1 和 2, 那么也有 $uwy \notin L$;
5. 与泵引理 $uwy = uv^0wx^0y \in L$ 矛盾, 假设不成立.
6. L 不是上下文无关的. □

例 2. 证明 $L=\{ww \mid w\in\{0,1\}^*\}$ 不是上下文无关的.

(错误的) 证明: 假设 L 是 CFL. 取 $z=0^N10^N1$, 那么
$z=uvwxy$ 为

$$z=\overbrace{\underbrace{00\cdots00}_{u}\ \underbrace{0}_{v}\ \underbrace{1}_{w}}^{0^N1}\ \overbrace{\underbrace{0}_{x}\ \underbrace{00\cdots01}_{y}}^{0^N1}$$

则对任意 $i\geq 0$, 有 $uv^iwx^iy\in L$, 满足泵引理. $\square$

(正确的) 证明: 假设 L 是 CFL. 取 $z=0^N1^N0^N1^N$, 将 z 分为 $z=uvwxy$ 时

❶ 若 vwx 在 z 中点的一侧, uv^0wx^0y 显然不可能属于 L;

❷ 若 vwx 包括 z 中点, 那么 uv^0wx^0y 为 $0^N1^i0^j1^N$, 也不可能属于 L.

所以假设不成立, L 不是 CFL. $\square$