

第二章作业 2

一、构造下列语言的**文法**，并判断文法的**类型**

(1) $\{w \mid w = w^T, w \in \{a,b,c\}^+\}$

解：

$S \rightarrow aSa \mid bSb \mid cSc \mid a \mid b \mid c \mid aa \mid bb \mid cc$

上下文无关文法 (线性文法)

注意：(a) $w \in \{a,b,c\}^+$ 而不是 $w \in \{a,b,c\}^*$

(b) 只考虑奇数串或偶数串

(c) 【共性问题】不终止问题，如仅一个 $S \rightarrow aSa$ 是没法生成句子的。

(2) $\{wxw^T \mid x, w \in \{a,b,c\}^+\}$

解：思路一：(中间部分记为X)

$S \rightarrow aSa \mid bSb \mid cSc \mid aXa \mid bXb \mid cXc$

$X \rightarrow a \mid b \mid c \mid aX \mid bX \mid cX$

上下文无关文法 (线性文法)

思路二：(只要首尾一个字母相同即可)

$S \rightarrow aXa \mid bXb \mid cXc$

$X \rightarrow a \mid b \mid c \mid aX \mid bX \mid cX$

上下文无关文法 (线性文法)

其可以改写为正则文法：

$S \rightarrow aA \mid bB \mid cC$

$A \rightarrow aA \mid bA \mid cA \mid aa \mid ba \mid ca$

$B \rightarrow aB \mid bB \mid cB \mid ab \mid bb \mid cb$

$C \rightarrow aC \mid bC \mid cC \mid ac \mid bc \mid cc$

注意：(a) X作为中间部分，不能作为S (换句话说， w 或 $w^T \in \{a,b,c\}^+$ ，不为空)

(3) $\{a^n b^k a^m \mid k=n+m, n,k,m \geq 0\}$

解：(重写为 $a^n b^n b^m a^m$)

$S \rightarrow AB$

$A \rightarrow aAb \mid \epsilon$

$B \rightarrow bBa \mid \epsilon$

上下文无关文法

注意：(a) 如果先考虑个数，那么位置可能也要移了 (不是可移可不移，除非句子都可以)

(4) $\{w \mid w \in \{a,b\}^+, \text{ 且 } w \text{ 中 } a \text{ 的个数是 } b \text{ 的个数的两倍}\}$

解：思路一：(先考虑 a 和 b 个数，再考虑 a 和 b 的可能位置，通过交换实现)

$S \rightarrow AAB \mid AABS$

$AB \rightarrow BA$

$BA \rightarrow AB$

$A \rightarrow a$

$B \rightarrow b$

上下文有关文法

思路二：(同时考虑 a 和 b 个数，以及 a、a 和 b 的可能位置)

$S \rightarrow AAB \mid ABA \mid BAA$ #最简单情况

$A \rightarrow AS \mid SA \mid a$ #任意 A 出现可以插入 S (AAB, ABA, BAA)

$B \rightarrow BS \mid SB \mid b$ #任意 B 出现可以插入 S (AAB, ABA, BAA)

上下文无关文法

(思考：为什么整体插入 S？是否存在分开插入 S【但不能解析为整体插入】的情况？测了 50 万个句子没找到反例)

错误思路：(递归定义)

$S \rightarrow aab \mid aba \mid baa$

$S \rightarrow aabS \mid abaS \mid baaS$

$S \rightarrow aaSb \mid abSa \mid baSa$

$S \rightarrow aSab \mid aSba \mid bSaa$

$S \rightarrow Saab \mid Saba \mid Sbaa$

原因是新加的 a、a、b 不仅可能插在 S 的前后位置，也可能插在 S 的中间位置，例如 a a a b b b a a a 和 a a a a b b a b b a a a。

思路三：(同时考虑 a 和 b 个数，以及 a、a 和 b 的可能位置)

$S \rightarrow WaWaWbW \mid WaWbWaW \mid WbWaWaW$

$W \rightarrow WaWaWbW \mid WaWbWaW \mid WbWaWaW \mid \epsilon$

上下文无关文法

(思考：是否存在结构为 $W_1aW_2aW_3bW_4$ 且 W_i 不一定满足个数条件的的句子，即不被该文法接收的句子？测了 50 万个句子没找到反例)

注意：(a) 没考虑位置或考虑不全

(b) 长度不为0或漏了6等。

(5) $\{w \in \{a,b\}^* \mid |w| \bmod 2 = 0, \text{ 且 } w \text{ 包括 } aba\}$ ，其中 $|w|$ 表示 w 的长度

解：

$S \rightarrow OabaE \mid EabaO$

$O \rightarrow aE \mid bE$

$E \rightarrow aO \mid bO \mid \epsilon$

上下文无关文法 (该语言是正则语言，正则文法留着思考题)

注意：(a) aba 的位置可以是前缀、后缀、(任意位置) 子串

(b) 奇数/偶数的前后缀可以互换

(c) 两个条件分开考虑！

(6) $\{ww : w \in \{a, b\}^*\}$ (选做)

解:

$S \rightarrow R \#$ (#表示最右边边界)

$R \rightarrow aRa$

$R \rightarrow bRb$ (这两条生成 ww^T)

$R \rightarrow T$ (T 表示左边部分的边界)

$T \rightarrow TQ$ (启动字符移动)

$Qaa \rightarrow aQa$

$Qab \rightarrow bQa$

$Qa\# \rightarrow \#a$ (以上三条将 Q 后面的 a 一直往右移动直到右边边界#)

$Qbb \rightarrow bQb$

$Qba \rightarrow aQb$

$Qb\# \rightarrow \#b$ (以上三条将 Q 后面的 b 一直往右移动直到右边边界#)

$T\# \rightarrow \varepsilon$ (没有字符需要移动, 则消去特殊边界字符)

无限制文法