

## Homework 2 — March 24

Lecturer: Mingji Xia

Completed by: 吉骏雄

## 第 2.1 次作业: 2.1(c), 2.4(b), 2.6(b), 2.14.

2.1 回忆一下例 2.3 中给出的 CFG  $G_4$ 。为方便起见, 用单字母重新命名它的变元如下:

$$E \rightarrow E + T \mid T$$

$$T \rightarrow T \times F \mid F$$

$$F \rightarrow (E) \mid a$$

给出下述字符串的语法分析树和派生:

c.  $a + a + a$

解 语法分析树如图 2.1 派生如下:

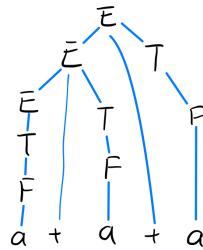


图 2.1. 2.1 c. 语法分析树

$$\begin{aligned} E &\Rightarrow E + T \Rightarrow E + T + T \Rightarrow T + T + T \Rightarrow F + T + T \Rightarrow a + T + T \\ &\Rightarrow a + F + T \Rightarrow a + a + T \Rightarrow a + a + F \Rightarrow a + a + a \end{aligned}$$

2.4 设字母表  $\Sigma$  是  $\{0, 1\}$ , 给出产生下述语言的上下文无关文法:

b.  $\{\omega \mid \omega \text{ 以相同的符号开始和结束}\}$

解 认为空串不属于这个语言的话, 那么文法应该是:

$$E \rightarrow 0F0 \mid 1F1 \mid 0 \mid 1$$

$$F \rightarrow 0F \mid 1F \mid \varepsilon$$

2.6 给出产生下述语言的上下文无关文法:

b. 语言  $\{a^n b^n \mid n \geq 0\}$  的补集

**解** 如果一个字符串  $\omega$  属于它的补集, 那么  $\omega$  一定满足四种形式之一:

1. 不包含  $b$ , 即  $a^m$  或空串  $\varepsilon$ .
2. 以  $b$  开头.
3. 以  $a^{r+s} b^s a$  开头, 或就是  $a^{r+s} b^s$ , 其中  $r, s > 0$ .
4. 以  $a^r b^{r+1}$  开头, 其后可能是空串, 也可能不是.

因此, 可以给出如下的上下文无关文法:

$$\begin{aligned} S &\rightarrow A \mid B \mid C \mid D \\ A &\rightarrow aA \mid \varepsilon \\ B &\rightarrow bR \\ C &\rightarrow aAGaR \mid aAG \\ D &\rightarrow GbR \\ G &\rightarrow aGb \mid ab \\ R &\rightarrow aR \mid bR \mid \varepsilon \end{aligned}$$

2.14 用定理 2.6 中给出的过程, 把下述 CFG 转换成等价的乔姆斯基范式文法。

$$\begin{aligned} A &\rightarrow BAB \mid B \mid \varepsilon \\ B &\rightarrow 00 \mid \varepsilon \end{aligned}$$

**解** 变化过程如下:

1. 添加新起始变元

$$\begin{aligned} S_0 &\rightarrow A \\ A &\rightarrow BAB \mid B \mid \varepsilon \\ B &\rightarrow 00 \mid \varepsilon \end{aligned}$$

2. 删除  $B \rightarrow \varepsilon$

$$\begin{aligned} S_0 &\rightarrow A \\ A &\rightarrow BAB \mid B \mid \varepsilon \mid AB \mid BA \mid A \\ B &\rightarrow 00 \end{aligned}$$

3. 删除  $A \rightarrow \varepsilon \mid A$

$$\begin{aligned} S_0 &\rightarrow A \mid \varepsilon \\ A &\rightarrow BAB \mid B \mid AB \mid BA \mid BB \\ B &\rightarrow 00 \end{aligned}$$

4. 删除  $A \rightarrow B$

$$\begin{aligned} S_0 &\rightarrow A \mid \varepsilon \\ A &\rightarrow BAB \mid AB \mid BA \mid BB \mid 00 \\ B &\rightarrow 00 \end{aligned}$$

5. 删除  $A \rightarrow BAB$

$$\begin{aligned} S_0 &\rightarrow A \mid \varepsilon \\ A &\rightarrow CB \mid AB \mid BA \mid BB \mid 00 \\ B &\rightarrow 00 \\ C &\rightarrow BA \end{aligned}$$

6. 删除  $S_0 \rightarrow A$

$$\begin{aligned} S_0 &\rightarrow CB \mid AB \mid BA \mid BB \mid 00 \mid \varepsilon \\ A &\rightarrow CB \mid AB \mid BA \mid BB \mid 00 \\ B &\rightarrow 00 \\ C &\rightarrow BA \end{aligned}$$

7. 删除  $S_0 \rightarrow 00, A \rightarrow 00, B \rightarrow 00$

$$\begin{aligned} S_0 &\rightarrow CB \mid AB \mid BA \mid BB \mid ZZ \mid \varepsilon \\ A &\rightarrow CB \mid AB \mid BA \mid BB \mid ZZ \\ B &\rightarrow ZZ \\ C &\rightarrow BA \\ Z &\rightarrow 0 \end{aligned}$$

**第 2.2 次作业: 2.11.**

**2.11** 用定理 2.12 中给出的过程, 把练习 2.1 中的  $\text{CFG}G_4$  转换成等价的 PDA.

**解** 结果如图 2.2

**第 2.3 次作业: 2.44, 2.2. 选做 2.37.**

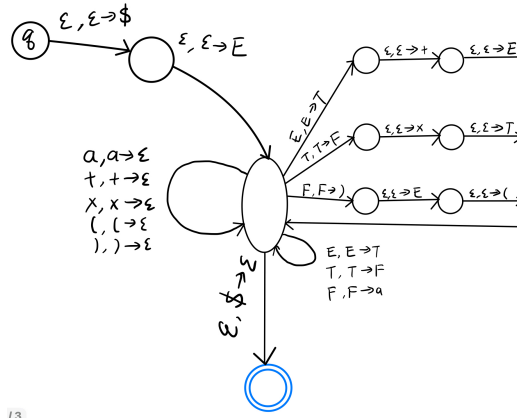


图 2.2. 2.11

**2.44** 对于字母表  $\Sigma = \{1, 2, 3, 4\}$  上的语言  $C = \{\omega \in \Sigma^* \mid \text{在 } \omega \text{ 中, } 1 \text{ 与 } 2 \text{ 的个数相同, } 3 \text{ 与 } 4 \text{ 的个数相同}\}$ , 证明  $C$  不是上下文无关语言.

**证明** 使用关于上下文无关语言的泵引理进行反证. 假设  $C$  是上下文无关语言, 那么  $\exists p \in \mathbb{Z}^+$ , 使得  $C$  中任何一个长度不小于  $p$  的字符串  $s$  都能被划分成 5 段  $s = uvxyz$ , 且满足下述条件:

1.  $\forall i \geq 0, uv^i xy^i z \in C$ ;
2.  $|vy| > 0$ ;
3.  $|vxy| \leq p$ .

由于语言的约束, 必须有  $vy$  中 1 与 2 的个数相等, 且 3 与 4 的个数相等. 构造字符串形式为  $1^{p+1}3^{p+1}2^{p+1}4^{p+1} \in C$ , 但无论如何使用分段方法与之匹配, 由于  $|vxy| \leq p$  的限制存在,  $vxy$  中至多存在 2 个不同的字符, 且一定是以下五种形式之一:  $1^m, 1^m 3^n, 3^m 2^n, 2^m 4^n, 4^m$ . 无论哪一种可能性, 作为其中的两个子串,  $v, y$  永远不可能满足  $vy$  中 1 与 2、3 与 4 的个数相等 (因为母串中不可能同时存在每一对中的两种字符!), 因此无法找出  $s = uvxyz$ , 矛盾! 所以  $C$  不是上下文无关语言.  $\square$

## 2.2

**a.** 利用语言  $A = \{a^m b^n c^n \mid m, n \geq 0\}$  和  $B = \{a^n b^n c^m \mid m, n \geq 0\}$  以及例 2.20, 证明上下文无关语言类在交运算下不封闭.

**证明** 考虑  $A$  与  $B$  的交集:  $A \cap B =: C = \{a^n b^n c^n \mid n \geq 0\}$ . 这很好理解, 因为  $A, B$  描述的均是  $a, b, c$  按顺序出现的字符串, 语言  $A$  中  $b, c$  的字数相同, 语言  $B$  中  $a, b$  的字数相同, 所以语言  $C = A \cap B$  中  $a, b, c$  的字数都相同. 接下来需要证明的是:  $A, B$  均为上下文无关语言, 而  $C$  不是. 证明使用关于上下文无关语言的泵引理.

$A = \{a^m b^n c^n \mid m, n \geq 0\}$  中任取字符串都有  $a^m b^n c^n$  的形式, 只需要取  $u = a^m b^{n-1}, v = b, x = \varepsilon, y = c, z = c^{n-1}$ , 这样  $uv^i xy^i z = a^m b^{n+i-1} c^{n+i-1} \in C$ . 这只需要  $p = 2$  即可.

$B = \{a^n b^n c^m \mid m, n \geq 0\}$  中任取字符串都有  $a^n b^n c^m$  的形式, 只需要取  $u = a^{n-1}, v = a, x = \varepsilon, y = b, z = b^{n-1} c^m$ , 这样  $uv^i xy^i z = a^{n+i-1} b^{n+i-1} c^m \in C$ . 这只需要  $p = 2$  即可.

$C = \{a^n b^n c^n \mid n \geq 0\}$  的情况则有不同. 假设其为上下文无关语言, 那么  $\exists p$  满足其条件; 这时取  $s = a^p b^p c^p$ , 发现无论如何取长度不超过  $p$  的子串, 总是只能包含  $\{a, b, c\}$  中的最多两个字符, 这样在将  $v, y$  乘幂叠加之后, 必然会产生  $a, b, c$  个数不相等的字符串 (比如  $uv^2xy^2z \notin C$ ). 所以  $C$  不是上下文无关语言.

$A \cap B \neq C$ , 因此上下文无关语言类在交运算下不是封闭的.  $\square$

b. 利用 (a) 和德·摩根律 (定理 0.10), 证明上下文无关语言类在补运算下不封闭.

**证明** 假设上下文无关语言类在补运算下封闭. 记上下文无关语言类为  $S$ , 已知  $S$  在并运算下封闭 (作为引理, 下面证明). 那么根据上一问的结果, 我们得到  $A, B \in S$  的交  $C \notin S$ . 因此有:  $C = A \cap B = \overline{\overline{A \cap B}} = \overline{\overline{A} \cup \overline{B}}$ . 然而,  $A, B \in S \implies \overline{A}, \overline{B} \in S \implies \overline{A} \cup \overline{B} \in S \implies C = \overline{\overline{A} \cup \overline{B}} \in S$ , 产生矛盾! 因此假设不成立, 上下文无关语言类在补运算下不封闭.

现在证明引理: 上下文无关语言类在并运算下封闭. 每一个 CFL 都有一个 CFG 可以生成它, 对任意的两个初始 CFL:  $A, B$ , 可以找出两个 CFG:  $G_A = \{V_A, \Sigma_A, R_A, S_A\}, G_B = \{V_B, \Sigma_B, R_B, S_B\}$  分别接受它们. 不妨令  $V_A \cap V_B = \emptyset, \Sigma_A \cap \Sigma_B = \emptyset$ . 构造一个新的 CFG:  $G_C = \{V_A \cup V_B \cup \{S_C\}, \Sigma_A \cup \Sigma_B, R_A \cup R_B \cup \{S_C \rightarrow S_A \mid S_B\}, S_C\}$ , 其中  $S_C \notin V_A \cup V_B$ . 这样, 由于新的起始变元一定会生成  $G_A, G_B$  两种文法的起始变元之一, 且规则没有任何重叠, 我们可以得知这个文法生成的语言一定是原本两种文法生成的语言的并集, 即  $G_C = G_A \cup G_B$ , 引理得证.  $\square$

**选做 2.37** 对任一语言  $A$ , 令运算符  $\text{SUFFIX}(A) = \{v \mid \text{存在串 } u, \text{使 } uv \in A\}$ , 证明上下文无关语言类在  $\text{SUFFIX}$  运算下封闭.

**证明** 对语言  $A$ , 可以找到一个满足乔姆斯基范式的上下文无关文法生成之, 记该文法为  $G_A = \{V, \Sigma, R, S\}$ . 为其中每一条形如  $V \rightarrow CD$  的规则进行对应的规则新增:  $V' \rightarrow C'D \mid CD \mid D \mid \omega$ , 并将初始变元  $S$  改为  $S'$ . 概括性的叙述, 即含“'”的变元 (对应原本存在的每一条规则) 可以进行上述四种情形的替换, 而原本的变元依旧只能进行原本的替换. 这样, 对于一个原本存在的派生树, 我们能在某个特殊的时间节点将之最左侧 (始终只存在一个的) 含“'”的新变元抹去, 产生的字符串的最左侧一段 (可以是  $\omega$ ) 也随之截去. 证明这样产生的新文法 (记作  $G'_A$ ) 可以产生的语言即为  $\text{SUFFIX}(A)$ . 方法即使用二叉派生树对字符串  $uv$  进行分解, 分两个方向进行证明.

(由于这是选做题, 就懒得写了, 以后有空的时候再打磨一下细节)  $\square$