第 6 次作业 – 上下文无关文法的性质

2154312 郑博远

习题 7.2.1 用 CFL 泵引理来证明下面的语言都不是上下文无关的：

a) $\{a^i b^j c^k \mid i < j < k\}$。

b) $\{a^n b^n c^i \mid i \le n\}$。

c) $\{0^p \mid p$ 是素数$\}$。提示：使用和例 4.3 中证明不是正则语言时采用相同的思想。

d) $\{0^i 1^j \mid j = i^2\}$。

e) $\{a^n b^n c^i \mid n \le i \le 2n\}$。

f) $\{ww^R w \mid w$ 是 0 和 1 的串$\}$。也就是说，由某个串 w 和它的反向串再和它本身连接起来的串（比如 001100001）构成的集合。

解答：

a) 假设该语言 L 是上下文无关的，且泵长度为 p。考察字符串 $s = a^p b^{p+1} c^{p+2} \in L$。设 $s = uvwxy$，由于$|vwx| \le p$，所以 vwx 不会同时含有三个字符。

1. vwx 只包含 a。由于 $|vx| \ge 1$，uv^2wx^2y中 a 的个数一定大于等于 b，因此 $uv^2wx^2y \notin L$；

2. vwx 只包含 b。由于 $|vx| \ge 1$，uv^2wx^2y中 b 的个数一定大于等于 c，因此 $uv^2wx^2y \notin L$；

3. vwx 只包含 c。由于 $|vx| \ge 1$，uv^0wx^0y中 c 的个数一定小于等于 b，因此 $uv^2wx^2y \notin L$；

4. v 中只含 a，x 中只含 b。由于 $|vx| \ge 1$，uv^2wx^2y中一定有 a 的个数大于等于 b（$|w| = 0$）或 b 的个数大于等于 c（$|w| \ne 0$），因此 $uv^2wx^2y \notin L$；

5. v 中只含 b，x 中只含 c。由于 $|vx| \ge 1$，若 $|v| = 0(|x| \ge 1)$，则 uv^0wx^0y中一定有 c 的个数小于等于 b，因此 $uv^0wx^0y \notin L$；若 $|v| \ne 0$，uv^0wx^0y中 b 的个数一定小于等于 a，因此 $uv^0wx^0y \notin L$；

6. v 或 x 中包含两种不同的符号。由于 $|vx| \ge 1$，uv^2wx^2y 中必将呈现两种字符交错出现的情况，因此$uv^2wx^2y \notin L$。

综上所述，L 不是 CFL。

b) 假设该语言 L 是上下文无关的，且泵长度为 p。考察字符串 $s=a^p b^p c^p \in L$。设 $s=uvwxy$，由于$|vwx| \le p$，所以 vwx 不会同时含有三个字符。

1. vwx 只包含 a。由于 $|vx| \ge 1$，uv^2wx^2y中 a、b个数不相等，$uv^2wx^2y \notin L$；

2. vwx 只包含 b。由于 $|vx| \ge 1$，uv^2wx^2y中 a、b个数不相等，$uv^2wx^2y \notin L$；

3. vwx 只包含 c。由于 $|vx| \ge 1$，uv^2wx^2y中 c 的个数一定大于 b，$uv^2wx^2y \notin L$；

4. v 中只含 a，x 中只含 b。由于 $|vx| \ge 1$，uv^0wx^0y 中 a 或 b 的个数一定小于 c，因此 $uv^0wx^0y \notin L$；

5. v 中只含 b，x 中只含 c。由于 $|vx| \ge 1$，若 $|v|=0$ $(|x| \ge 1)$，则 uv^2wx^2y中 c 个数一定大于 b；若 $|v| \ne 0$，uv^2wx^2y中 a 与 b 的个数一定不相等。因此 $uv^2wx^2y \notin L$；

6. v 或 x 中包含两种不同的符号。由于 $|vx| \ge 1$，uv^2wx^2y 中必将呈现两种字符交错出现的情况，因此 $uv^2wx^2y \notin L$。

综上所述，L 不是 CFL。

c) 假设该语言 L 是上下文无关的，且泵长度为 p。考察字符串 $s=0^{p_0}$（p_0是大于p的素数）$\in L$，则 $s'=uv^kwx^ky=0^{p_0+(k-1)|vx|} \in L$。取 $k=p_0+1$，则 $s'=0^{p_0(|vx|+1)}$。由于 $|vx| \ge 1$，$p_0(|vx|+1)$ 不是素数，因此 $uv^kwx^ky \notin L$，与假设矛盾。

综上所述，L 不是 CFL。

d) 假设该语言 L 是上下文无关的，且泵长度为 p。考察字符串 $s=0^p1^{p^2} \in L$。设 $s=uvwxy$。

1. vwx 只包含 0。由于 $|vx| \ge 1$，显然 uv^2wx^2y中 1 的个数与 0 的个数不满足平方关系，$uv^2wx^2y \notin L$；

2. vwx 只包含 1。由于 $|vx| \ge 1$，显然 $uv^2\ wx^2\ y$中 1 的个数与 0 的个数不满足平方关系，$uv^2wx^2y \notin L$；

3. v 中只含 0，x 中只含 1。考察串 $uv^2wx^2y=0^{p+|v|}1^{p^2+|x|}$。$(p+|v|)^2=p^2+2p|v|+|v|^2$。当 $|v|=0$ 时，由于 $|vx| \ge 1$ 有 $|x| \ge 1$，则 $(p+|v|)^2=p^2 \ne p^2+|x|$；当 $|v| \ne 0$ 时，由于 $|x| \le p$，$(p+|v|)^2 < p^2+x$。因此，$(p+|v|)^2 \ne p^2+x$，$uv^2wx^2y \notin L$。

4. v 或 x 中包含两种不同的符号。由于 $|vx| \ge 1$，uv^2wx^2y 中必将呈现两种字符交错出现的情况，因此 $uv^2wx^2y \notin L$。

综上所述，L 不是 CFL。

e) 假设该语言 L 是上下文无关的，且泵长度为 p。考察字符串 $s=a^p b^p c^p \in L$。设

$s = uvwxy$，由于$|vwx| \le p$，所以 vwx 不会同时含有三个字符。

1. vwx 只包含 a。由于 $|vx| \ge 1$，uv^2wx^2y中 a、b个数不相等，$uv^2wx^2y \notin L$；

2. vwx 只包含 b。由于 $|vx| \ge 1$，uv^2wx^2y中 a、b个数不相等，$uv^2wx^2y \notin L$；

3. vwx 只包含 c。由于 $|vx| \ge 1$，uv^0wx^0y中 c 的个数一定小于 b，$uv^2wx^2y \notin L$；

4. v 中只含 a，x 中只含 b。由于 $|vx| \ge 1$，uv^2wx^2y 中 a 或 b 的个数一定大于 c，因此 $uv^2wx^2y \notin L$；

5. v 中只含 b，x 中只含 c。由于 $|vx| \ge 1$，若 $|v| = 0$ ($|x| \ge 1$)，则 uv^0wx^0y中 b 的个数一定大于 c 的个数，因此 $uv^0wx^0y \notin L$。若 $|v| \ne 0$，uv^0wx^0y中 a 与 b 的个数一定不相等，因此 $uv^0wx^0y \notin L$；

6. v 或 x 中包含两种不同的符号。由于 $|vx| \ge 1$，uv^2wx^2y 中必将呈现两种字符交错出现的情况，因此 $uv^2wx^2y \notin L$。

综上所述，L 不是 CFL。

f) 假设该语言 L 是上下文无关的，且泵长度为 p。考察字符串 $s = 0^p1^{2p}0^{2p}1^p \in L$。设 $s = uvwxy$，由于$|vwx| \le p$，所以 vwx 不会横跨 3 个 0、1 段。

1. v 和 x 都在第一部分的 0 中。由于 $|vx| \ge 1$，$uv^2wx^2y = 0^{p+|vx|}1^{2p}0^{2p}1^p$ 无法被划分为 ww^Rw 的形式（$2p \ne 2(p+|vx|)$，无法划分成 w^Rw），$uv^2wx^2y \notin L$；

2. v 和 x 都在第一部分的 1 中。由于 $|vx| \ge 1$，同理 $uv^2wx^2y = 0^p1^{2p+|vx|}0^{2p}1^p$ 无法被划分为 ww^Rw 的形式，$uv^2wx^2y \notin L$；

3. v 和 x 都在第二部分的 0 中。由于 $|vx| \ge 1$，同理 $uv^2wx^2y = 0^p1^{2p}0^{2p+|vx|}1^p$ 无法被划分为 ww^Rw 的形式，$uv^2wx^2y \notin L$；

4. v 和 x 都在第二部分的 1 中。由于 $|vx| \ge 1$，同理 $uv^2wx^2y = 0^p1^{2p}0^{2p}1^{p+|vx|}$无法被划分为 ww^Rw 的形式，$uv^2wx^2y \notin L$；

5. v 在第一部分的 0 中，x 在第一部分的 1 中。由于 $|vx| \ge 1$，同理 $uv^2wx^2y = 0^{p+|v|}1^{2p+|x|}0^{2p}1^p$无法被划分为 ww^Rw 的形式，$uv^2wx^2y \notin L$；

6. v 在第一部分的 1 中，x 在第二部分的 0 中。由于 $|vx| \ge 1$，同理 $uv^2wx^2y = 0^p1^{2p+|v|}0^{2p+|x|}1^p$无法被划分为 ww^Rw 的形式，$uv^2wx^2y \notin L$；

7. v 在第二部分的 0 中，x 在第二部分的 1 中。由于 $|vx| \ge 1$，同理 $uv^2wx^2y = 0^p1^{2p}0^{2p+|v|}1^{p+|x|}$无法被划分为 ww^Rw 的形式，$uv^2wx^2y \notin L$；

8. v 或 x 中包含两种不同的符号。打圈的部分会产生小于 p 长度的 0、1 字符交错，

但其他部分的 0、1 部分长度均大于 p，因此无法找到对应的另外两部分串。即 $uv^2wx^2y=$ 无法被划分为 ww^Rw 的形式，$uv^2wx^2y \notin L$ 。

综上所述，L 不是 CFL。

习题 7.2.5 使用奥格登引理（习题 7.2.3）来证明下列语言不是 CFL：

a) $\{0^i1^j0^k \mid j = \max(i,k)\}$。

b) $\{a^nb^nc^i \mid i \neq n\}$。提示：如果 n 是奥格登引理的常数，考虑串 $z = a^nb^nc^{n+n!}$。

解答：

a) 假设该语言 L 是上下文无关的，且泵长度为 p。考察字符串 $s=0^{2p}1^{2p}0^p \in L$，选择末尾部分的 n 个 0 为显著位置。设 $s=uvwxy$，由奥格登引理可得 vx 中至少有一个显著位置。因此，x 中一定包含末尾部分的 0。

1. v 或 x 中包含两种不同的符号。由于 $|vx| \geq 1$，uv^2wx^2y 中必将呈现两种字符交错出现的情况，因此 $uv^2wx^2y \notin L$；

2. v 中只含开始部分的 0，x 中只含末尾部分的 0。由于 vx 中至少有一个显著位置，$|x| \geq 1$。考察字符串$uv^{2p+1}wx^{2p+1}y$，末尾部分的 0 个数一定大于 $2p$，但 1 的个数仍然为 $2p$，不满足 $j=\max(i,k)$，因此 $uv^2wx^2y \notin L$；

3. v 中只含 1，x 中只含末尾部分的 0。考虑将 v 和 x 打 i 圈的情况，则 $s'=0^{2p}1^{2p+(i-1)|v|}0^{p+(i-1)|x|}$。显然有 i 满足 $p+(i-1)|x|>2p$，且 $2p+(i-1)|v| \neq p+(i-1)|x|$，此时 $uv^iwx^iy \notin L$。

4. v 和 x 中都只含末尾部分的 0。由于 vx 中至少有一个显著位置，考察字符串 $uv^{2p+1}wx^{2p+1}y$，末尾部分的 0 个数一定大于 $2p$，但 1 的个数仍然为 $2p$，不满足 $j=\max(i,k)$，因此 $uv^2wx^2y \notin L$。

综上所述，L 不是 CFL。

b) 假设该语言 L 是上下文无关的，且泵长度为 n。考察字符串 $s=a^nb^nc^{n+n!} \in L$，选择所有的字符 a 为显著位置。设 $s=uvwxy$，由奥格登引理可得 vx 中至少有一个显著位置。因此，v 中一定包含字符 a。

1. v 或 x 中包含两种不同的符号。由于 $|vx| \geq 1$，uv^2wx^2y 中必将呈现两种字符交

错出现的情况，因此 $uv^2wx^2y \notin L$；

2. v 中仅含 a，x 中仅含 a。由于 $|vx| \geq 1$，uv^2wx^2y 中必然有 a 的字符个数大于 b 的字符个数，因此 $uv^2wx^2y \notin L$；

3. v 中仅含 a，x 中仅含 b。若 $|v| \neq |x|$，uv^2wx^2y 中必然有 a 的字符个数与 b 不相等，因此 $uv^2wx^2y \notin L$。若 $|v| = |x|$，打 i 圈后有 $uv^iwx^iy = a^{n+(i-1)|v|}b^{n+(i-1)|v|}c^{n+n!}$。由于 $|v| \leq n$，取 $i = \frac{n!}{|v|} + 1$。此时，$uv^iwx^iy = a^{n+n!}b^{n+n!}c^{n+n!} \notin L$；

4. v 中仅含 a，x 中仅含 c。由于 vx 中至少有一个显著位置，v 中一定包含字符 a，uv^2wx^2y 中必然有 a 的字符个数大于 b 的字符个数，因此 $uv^2wx^2y \notin L$。

综上所述，L 不是 CFL。

补充习题 1 构造与下列文法等价的 CNF。

$$S \to ABB \mid bAA$$

$$B \to aBa \mid aa \mid \varepsilon$$

$$A \to bbA \mid \varepsilon$$

解答：

消除 $\varepsilon-$产生式：

$$S \to ABB \mid BB \mid AB \mid A \mid B \mid bAA \mid bA \mid b \mid \varepsilon$$

$$B \to aBa \mid aa$$

$$A \to bbA \mid bb$$

消除单一产生式：

$$S \to ABB \mid BB \mid AB \mid bAA \mid bA \mid b \mid \varepsilon \mid aBa \mid aa \mid bbA \mid bb$$

$$B \to aBa \mid aa$$

$$A \to bbA \mid bb$$

引入新变元：

$$S \to AC \mid BB \mid AB \mid G_bD \mid G_bA \mid G_aE \mid G_aG_a \mid FA \mid G_bG_b \mid b \mid \varepsilon$$

$$B \to G_aE \mid G_aG_a$$

$$A \to FA \mid G_bG_b$$

$$G_a \rightarrow a$$

$$G_b \rightarrow b$$

$$C \rightarrow BB$$

$$D \rightarrow AA$$

$$E \rightarrow BG_a$$

$$F \rightarrow G_bG_b$$