形式语言与自动机 作业三

cycleke

1 第一题

Design context-free grammars for the following languages:

1.1 (a)

The set $\{a^i b^j c^k | i \neq j \ or \ j \neq k\}$, that is, the set of strings of a's followed by b's followed by c's, such that there are either a different number of a's and b's or a different number of b's and c's, or both.

解 1.1 我们定义一个 CFG $G = (\{S, A, B, C, D, E\}, \{a, b, c\}, P, S)$，其中 P 为

$$
\begin{aligned}
S &\rightarrow ADC|DBC|ABE|AEC|AD|DB|BE|EC \\
A &\rightarrow aA|a \\
B &\rightarrow bB|b \\
C &\rightarrow cC|c \\
D &\rightarrow aDb|\varepsilon \\
E &\rightarrow bEc|\varepsilon
\end{aligned}
$$

其中 A, B, C 表示各个字符的正闭包，D 和 E 则分别表示 ab 相等和 bc 相等的字符串，而我们按照 i, j, k 的关系来生成 S。

1.2 (b)

The set of all strings a's and b's that are not of the form ww, that is, not equal to any string repeated.

解 1.2 我们定义一个 CFG $G = (\{S, A, B\}, \{a, b\}, P, S)$，其中 P 为

$$
\begin{aligned}
S &\rightarrow AB|BA|A|B \\
A &\rightarrow aAa|aAb|bAa|bAb|a \\
B &\rightarrow aBa|aBb|bBa|bBb|b
\end{aligned}
$$

其中 A, B 表示中心为 a, b 的奇数长度的字符串，于是我们可以使用 A, B 来表示奇数长度的串而对于偶数长度的串，我们利用 AB, BA 来构造，而 AB 的中心不同所以他们的连接一定不是回文串。

1.3 (c)

The set of all strings with twice as many 0's as 1's.

解 1.3 我们定义一个 CFG $G = (\{S\}, \{0, 1\}, P, S)$，其中 P 为

$$S \to \varepsilon | 1S00 | 00S1 | 0S1S0 | SS$$

1.4 (d)

$L(00^*11^*22^*00^*11^*22^*00^*11^*22^*)$

解 1.4 我们定义一个 CFG $G = (\{S\}, \{0, 1, 2\}, P, S)$，其中 P 为

$$\begin{aligned}
S &\to AAA \\
A &\to BCD \\
B &\to 0|0B \\
C &\to 1|1C \\
D &\to 2|2C
\end{aligned}$$

2 第二题

Consider the CFG G defined by productions:

$$S \to aS|Sb|a|b$$

Prove by induction on the string length that no string in L(G) has ba as a substring.

证明 我们需要证明 $w \in L(G)$，w 不包含子串 ba。

基础：当 $|w| = 1$ 时，显然 w 不包含子串 ba。

归纳：假设对于所有的 $n \leq k, w \in L(G), |w| = n$，串 w 均不包含子串 ba 。对于串 $w, |w| = k + 1$，根据 G 的文法，我们可以分为两种情况

- 由 $S \to aS$，有 $w = as, |s| = k$。依据归纳假设，s 不包含子串 ba。我们只是在 s 头部添加了一个字符 a，只会产生两种新子串 aa, ab，而 $aa \neq ba, ab \neq ba$ 所以不会产生子串 ba。
- 由 $S \to Sb$，有 $w = sb, |s| = k$。依据归纳假设，s 不包含子串 ba。我们只是在 s 尾部添加了一个字符 b，只会产生两种新子串 ab, bb，而 $ab \neq ba, bb \neq ba$ 所以不会产生子串 ba。

所以对于长度为 $k + 1$ 的串，它亦不包含子串 ba。

综上所述，L(G) 中的字符串不会包含 ba。 □