arXiv:2603.27656v1 [cs.IT] 29 Mar 2026

有限码中交换等价的一种弱结构形式

Dean Kraizberg*

2026年3月31日

摘要

我们研究二元字母表上的前缀无关码与一类无标号根树之间的结构关系，并将这类树称为对称树。我们建立了前缀无关码与对称树之间的规范对应关系，该对应关系不仅保留码字的长度，还保留某些额外的交换结构。利用这一对应关系，我们给出了一个与交换等价猜想相关的结果。我们证明，对于任意码，都存在一个前缀无关码，使得对于每个固定的词长，由指定符号出现次数所确定的 2 的幂之和相等。

关键词： 前缀无关码、对称树、交换等价、自由幺半群。*MSC2020*： 68R15、20M05、94A45。

1 引言

1.1 关于码的一些背景

关于编码理论的基础介绍，我们请读者参阅 [5]。我们只介绍必要的符号和定义。全文中，我们都在二元字母表$\{a,b\}$上进行讨论。

记号 1.1 我们用 $\{a,b\}^{<\mathbb{N}}$ 表示由 $\{a,b\}$ 生成的自由幺半群；也就是说，它是 $\{a,b\}$ 上所有有限词构成的集合：

$$\{a,b\}^{<\mathbb{N}} = \{\langle x_1,\dots,x_n\rangle : n \geq 1,\ x_i \in \{a,b\}\} \cup \{\langle\rangle\},$$

其中，$\langle\rangle$ 表示空词。

必要强调词的序列结构时，我们用尖括号表示词。对于一个词 $w = \langle x_1,\dots, x_n\rangle \in \{a,b\}^{<\mathbb{N}}$，我们将其长度定义为$|w| = n$。

给定两个字 $w_1 = \langle x_1,\dots,x_n\rangle$ 和 $w_2 = \langle y_1,\dots,y_m\rangle$，它们的连接定义为

$$w_1 w_2 = \langle x_1,\dots,x_n,y_1,\dots,y_m\rangle.$$

*特拉维夫大学数学科学学院，deank@mail.tau.ac.il

现在我们回顾一下（唯一可译的有限）码的概念。从这里开始，我们将这种对象简称为码。

定义 1.2 码 是由有限词构成的有限集 $C \subseteq \{a,b\}^{<\mathbb{N}}$，并满足以下性质：对于任意

$$w_1 \cdots w_n = w'_1 \cdots w'_m,$$

$with$ $w_1, \ldots, w_n, w'_1, \ldots, w'_m \in C$，则对所有 $1 \leq i \leq n$， $n = m$ 和 $w_i = w'_i$。

克拉夫特–麦克米伦定理给出了码字长度的一项基本约束。该定理不仅断言任何码中的码字长度都必须满足某个约束，还指出每个满足此条件的序列，都可以由某个前缀无关码的码字长度构成。

定理 1。 3（克拉夫特–麦克米伦）设 C 是一个定义在大小为 k 的字母表上的码。n

$$\sum_{w \in C} k^{-|w|} \leq 1.$$

反过来，设 A 是一个满足以下条件的正整数有限多重集

$$\sum_{a \in A} k^{-a} \leq 1.$$

那么，存在一个前缀无关码 C，使得 C的码字长度多重集恰为 A，也就是说，$\{|c| : \ c \in C\} = A$ 作为多重集。

特别地，每个码都与一个前缀无关码长度等价；也就是说，对于每个码 C，都存在一个前缀无关码 C' 以及一个双射 $f: \ C \to C'$，使得对所有 $w \in C$，$|w| = |f(w)|$。

Perrin 和 Sch¨utzenberger 引入了一种更强的等价概念 [3]。对于一个字 $w = \langle x_1, \ldots, x_n \rangle \in \{a,b\}^n$，我们定义

$$[w]_a := \#\{1 \leq i \leq n : x_i = a\}, \qquad [w]_b := \#\{1 \leq i \leq n : x_i = b\}.$$

定义 1.4 若存在一个双射 $f: \ C_1 \to C_2$，使得对每个 $w \in C_1$，则称两个码 C_1、C_2 是交换等价 的。

$$[w]_a = [f(w)]_a \quad and \quad [w]_b = [f(w)]_b.$$

最初，人们猜想每个码都与某个前缀无关码 [3]交换等价。然而，Peter Shor [7], 找到了一个码，推翻了这一猜想。

$$C_{\text{Shor}} = b\{1, a, a^7, a^{13}, a^{14}\} \cup \{a^3, a^8\}b\{1, a^2, a^4, a^6\} \cup a^{11}b\{1, a, a^2\},$$

它与任何前缀无关码都不交换等价。另见 [1]，其中有更多反例。鉴于这个反例，猜想被限制在一类特殊码类中 [6],

猜想 1.5（Sch¨utzenberger） 每个极大码——也就是说，对于每个 $w \in \{a,b\}^{<\mathbb{N}} \setminus C$，集合 $C \cup \{w\}$ 都不是码的码 C——都与某个前缀无关码交换等价。

上述猜想也称为交换等价猜想。关于交换等价猜想已有一些部分结果，例如见 [3, 9]。我们的主要结果——定理1.9——与该猜想有一定关联，因为它证明了存在一种前缀无关码，能够保持与码中符号出现次数相关的某些量。

1.2 对称树与主要结果

现在我们介绍研究树所采用的框架。这种表述符合有根树的标准图论概念。有关图论的基本术语和定义，请读者参见 [8] 。

定义 1.6 有根树是一个有限、连通的无环图 T，并带有一个指定的顶点 $o \in V(T)$，称为根。我们用dist 表示 T上的通常图距离。

给定顶点 u、$v \in V(T)$，若 u和 v相邻，且 u是 v的子节点，则称

$$\mathrm{dist}(u,o) = \mathrm{dist}(v,o) + 1.$$

对于 $v \in V(T)\setminus\{o\}$，我们用 T_v表示以 v为根的子树，它由 v及其所有后代组成。

叶节点是没有子节点的顶点——等价地，v 是叶节点，当且仅当 $T_v = \{v\}$是只有一个顶点的树。对于 $k \in \mathbb{N}$，我们定义

$$\mathcal{L}_k(T) := \{v \in V(T) : v \textit{ is a leaf and } \mathrm{dist}(v,o) = k\}.$$

接下来，我们引入有根树的一个对称性条件。

定义 1.7 设 T 是以 o为根的有限有根树，并假设每个顶点至多有三个子节点（也就是说，T 是完全3叉树的一棵有限子树）。

如果对每个至少有两个子节点的顶点 $v \in V(T)$，都存在 v 的两个不同子节点 u_1、u_2，使得有根子树 T_{u_1} 和 T_{u_2} 完全相同，则称 T是对称的。

现在我们陈述主要结果，这些结果说明了在研究前缀无关码时引入对称树的理由。

定理 1.8 前缀无歧义码与对称树之间存在双射对应。此外，在此对应关系下，若 C 是一个前缀无关码，且 T是其对应的对称树，那么对每个 $k \in \mathbb{N}$，

$$|\mathcal{L}_k(T)| = \sum_{\substack{w \in C \\ |w|=k}} 2^{[w]_a}.$$

我们将利用这一对应关系来证明以下结果：

定理 1.9 对于每个码 C，都存在一个前缀无关码 C'，使得对于每个$k \in \mathbb{N}$，都有

$$\sum_{\substack{w\in C\\|w|=k}} 2^{[w]_a} = \sum_{\substack{w'\in C'\\|w'|=k}} 2^{[w']_a}.$$

注 1.10 一般而言，定理 C' 中的码 *1.9* 不一定与 C具有相同的基数。事实上，*Shor* 的反例表明，这种情况一般并不成立。

为说明这一点，假设存在一个前缀无关码 C 满足定理 *1.9* 的结论，且 $|C| = |C_{\text{Shor}}|$。那么 C 与 C_{Shor} 必然交换等价，这与 *Shor* 的结果矛盾。

2 主要结果的证明

2.1 定理 1.8 的证明

我们先证明第一个方向，其思路与 [2] 中的论证类似。为此，我们引入若干辅助概念。

我们首先用有限词给出有根标记树的一种等价表述。该表述与 [4]中给出的表述一致。

定义 2.1 设 $\mathcal{A}$ 为有限字母表。定义在 $\mathcal{A}$ 上的标记树 是有限词构成的非空集合 $T \subseteq \mathcal{A}^{<\mathbb{N}}$，且 T 是前缀闭的；也就是说，若 $w \in T$ 且 w'是 w的前缀，则 $w' \in T$。

T中的元素称为顶点[1]。给定 w、$w' \in T$，若 w'恰好比 w多一个符号，即 w' 恰好在 w 后扩展一个符号，则称 w' 是 w的子节点，也就是说，

$$|w'| = |w| + 1 \quad and \quad w \ is \ a \ prefix \ of \ w'.$$

对于 $w \in T$，我们将以 w 为根的子树 定义为

$$T_w := \{\, w' \in T : w \ is \ a \ prefix \ of \ w' \,\}.$$

不难验证，该定义与通常图论意义上的有根标记树概念等价。

接下来，我们介绍自由群及其对应的凯莱图。

定义 2.2 设 $\mathcal{A}$ 是一个有限集。自由群 $F(\mathcal{A})$ 是由字母表 $\mathcal{A} \cup \mathcal{A}^{-1}$上的所有约化字构成的群，其中约化是指没有任何符号与其逆元相邻出现。

在 $\mathcal{A} = \{a, b\}$的情况下，我们记作 $F(\mathcal{A}) = F(a, b)$。

定义 2.3 自由群 $F(\mathcal{A})$的凯莱图，记作 $\mathrm{Cay}(F(\mathcal{A}))$，定义如下：

[1]全文中，顶点和码字一样，均用尖括号 $\langle\cdot\rangle$表示。

1. 顶点集为 $F(\mathcal{A})$。

2. 对于每个 $w \in F(\mathcal{A})$ 以及每个 $s \in \mathcal{A} \cup \mathcal{A}^{-1}$，存在一条标记为 s 的边，连接w 和 ws。

现在我们完成证明。

设 C 是字母表 $\{a, b\}$上的一个前缀无关码。定义

$$Z_b(C) := \{\langle b, c_1, b, c_2, \ldots, b, c_n\rangle : \langle c_1, \ldots, c_n\rangle \in C\},$$

并令 T_b 表示与之关联的标记树，其叶节点恰好对应于 C的元素。

考虑映射

$$\pi \colon \mathrm{Cay}(F(a,b)) \to \{a,b\}^{<\mathbb{N}}, \qquad \pi\big(\langle c_1^{\varepsilon_1}, \ldots, c_n^{\varepsilon_n}\rangle\big) = \langle c_1, \ldots, c_n\rangle,$$

其中 $c_i \in \{a, b\}$ 且对所有 i， $\varepsilon_i \in \{\pm 1\}$ 。

我们考虑凯莱图 Cay($\pi^{-1}(T_b)$ 中的提升子树$F(a, b)$。文献 [2] 证明，由那些字在每个奇数位置都含有符号 b 的顶点组成的诱导子树，对应于一个定义在大小为 3 的字母表上的前缀无关码。

令 T 为将此子树限制在其偶数坐标上所得的树。更明确地说， T由自由字母表（形式上） $\{a, a^{-1}, b\}$上的所有字 $\langle c'_1, \ldots, c'_n\rangle$ 组成，使得

$$\langle b, c'_1, b, c'_2, \ldots, b, c'_n\rangle \in \pi^{-1}(T_b).$$

我们声称，底层（无标记）树 T 是对称的。

事实上，设 $w \in T$ 是一个至少有两个子节点的顶点。那么存在不同的元素 s_1, $s_2 \in \{a, b\} \cup \{a^{-1}, b^{-1}\}$，使得 $ws_1, ws_2 \in T$。我们首先注意到，这些扩展都不可能涉及 b^{-1}，因为这会在 $\pi^{-1}(T_b)$ 中产生一个包含形如 bb^{-1} 的子字的约化字，这是不可能的。

此外， s_1， s_2 中至多只有一个可以等于 b。因此，不失一般性，我们可以假设 $s_1 = a$。根据构造的对称性，同样可得 $wa^{-1} \in T$ 。

由此可知，以 wa 和 wa^{-1} 为根的子树（忽略边标记）是同构的，因此 T 满足对称性条件。

最后，文献 [2] 证明了， T 中经 π 映射到某个给定叶节点的叶节点数量

$$\langle b, c_1, \ldots, b, c_k\rangle \in T_b, \qquad \text{where } \langle c_1, \ldots, c_k\rangle = w \in C,$$

恰好为

$$2^{|\{i \in [k] : c_i \neq b\}|} = 2^{[w]_a}.$$

因此，

$$|\mathcal{L}_k(T)| = \sum_{\substack{w \in C \\ |w| = k}} 2^{[w]_a},$$

如所要求。

现在我们证明反向结论。

设 T 是一棵对称有根树。我们用自由幺半群 $\{a, a^{-1}, b\}^{<\mathbb{N}}$ 中的元素为 T 构造一个标记。

我们进行归纳构造。给根 o 赋予空词，记作 $\langle \qquad\qquad\qquad\qquad \rangle$.

假设顶点 $v \in V(T)$ 已被赋予标记 $w_v \in \{a, a^{-1}, b\}^{<\mathbb{N}}$。下面说明如何标记它的子节点。

- 如果 v 有三个子节点 u_1, u_2, u_3，且其中两个（比如 u_1, u_2）满足 $T_{u_1} \cong T_{u_2}$，则赋予
$$w_{u_1} := w_v a, \qquad w_{u_2} := w_v a^{-1}, \qquad w_{u_3} := w_v b.$$
- 如果 v 恰有两个子节点 u_1, u_2，且满足 $T_{u_1} \cong T_{u_2}$，则赋予
$$w_{u_1} := w_v a, \qquad w_{u_2} := w_v a^{-1}.$$
- 如果 v 只有一个子节点 u，则赋予
$$w_u := w_v b.$$

这个标记定义了从 T 到自由幺半群 $\{a, a^{-1}, b\}^{<\mathbb{N}}$ 的嵌入。略微滥用记号，我们仍将其记作 T。

现在考虑投影映射

$$\mathcal{P}\colon \{a, a^{-1}, b\}^{<\mathbb{N}} \to \{a, b\}^{<\mathbb{N}}, \qquad \mathcal{P}\big(\langle a\rangle\big) = \mathcal{P}\big(\langle a^{-1}\rangle\big) = a,\ \mathcal{P}\big(\langle b\rangle\big) = b,\ \mathcal{P}(w_1 \cdot w_2) = \mathcal{P}(w_1)\mathcal{P}(w_2)$$

容易验证， $\mathcal{P}(T)$ 是一棵标记树，其叶节点集合构成一个前缀无关码 C，使得

$$|\mathcal{L}_k(T)| = \sum_{\substack{w \in C \\ |w| = k}} 2^{[w]_a},$$

如所需。 ■

注记 2.4 上述构造特别表明，对称树存在一种以 $F(a, b)$ 中元素为标记的规范标记，因此确定了唯一一个与之关联的前缀无关码。

2.2 定理证明 1.9

我们先陈述两个辅助引理。

引理 2.5 设 C 是一个码。则

$$\sum_{c \in C} 3^{-|c|}\, 2^{[c]_a} \le 1.$$

证明。 使用定理证明中的记号1.8，定义

$$\hat{C} = \left\{ \langle c'_1, \dots, c'_n \rangle : \langle b, c'_1, b, c'_2, \dots, b, c'_n \rangle \in \pi^{-1}(T_b) \right\} \subseteq \{a, a^{-1}, b\}^{<\mathbb{N}}.$$

ˆ 关键观察是， C 构成一个大小为 3 的字母表上的码。

ˆ 事实上，假设一个字有两种不同的分解，且分解因子均属于 C:

$$w_1 \cdots w_n = w'_1 \cdots w'_m, \qquad w_1, \dots, w_n, w'_1, \dots, w'_m \in \hat{C}.$$

对两边应用投影 $\mathcal{P}$，得到同一个字的两种分解，且分解因子均属于 C。由于 C 是一个码，因此

$$n = m \quad \text{and} \quad \mathcal{P}(w_i) = \mathcal{P}(w'_i) \quad \text{for all } 1 \le i \le n.$$

由于 $w_1 \cdots w_n = w'_1 \cdots w'_n$，我们得出，对于每个 $1 \le i \le n$，都有 $w_i = w'_i$，这与ˆ两种分解互不相同的假设矛盾。因此， C 是一个码。

ˆ 现在，将麦克米伦定理应用于 C（将其视为一个以 3 个字母组成的字母表为基础的码），再结合 [2]中定理 3 的证明所用的相同计数论证，即可得到所需的不等式。 ■

引理 2.6 设 A 是一个由自然数组成的有限多重集，并令 $N \in \mathbb{N}$。假设

$$\sum_{n \in A} 2^n \ge 2^N \quad \textit{and} \quad n \le N \textit{ for all } n \in A.$$

那么存在一个子多重集 $A' \subseteq A$，使得 $\sum_{n \in A'} 2^n = 2^N$。

证明。 我们对 N 使用归纳法。

基例： $N = 0$。在这种情况下， A 中唯一可能的元素是 0，而条件 $\sum_{n \in A} 2^n \ge 2^0 = 1$ 意味着 A 至少包含一个 0。令 A' 为从 A 中取出的恰好含一个 0 的多重集。那么 $\sum_{n \in A'} 2^n = 2^0 = 1$, 如所需。

归纳步骤： 考虑 $N > 0$，并假设该命题对所有小于 N 的自然数都成立。

如果 $N \in A$，那么我们可以取 $A' = \{N\}$，结论立即成立。否则， $N \notin A$。根据假设， $\sum_{n \in A} 2^n \ge 2^N \ge 2^{N-1}$。将归纳假设应用于 $N - 1$，可知存在一个子多重集 $A_1 \subseteq A$，使得 $\sum_{n \in A_1} 2^n = 2^{N-1}$.。考虑剩余的多重集 $A \setminus A_1$。我们有

$$\sum_{n \in A \setminus A_1} 2^n = \sum_{n \in A} 2^n - \sum_{n \in A_1} 2^n \ge 2^N - 2^{N-1} = 2^{N-1}.$$

再次对 $N - 1$ 应用归纳假设于 $A \setminus A_1$，我们得到一个子多重集 $A_2 \subseteq A \setminus A_1$，使得 $\sum_{n \in A_2} 2^n = 2^{N-1}$.。令 $A' = A_1 \cup A_2$，则有 $\sum_{n \in A'} 2^n = 2^N$，正如所需。 ■

现在我们转而证明主要结果。设 C 是一个码。根据定理1.8，只需构造一棵对称树 T 使得

$$|\mathcal{L}_k(T)| = \sum_{\substack{w \in C \\ |w| = k}} 2^{[w]_a}.$$

我们对 $|C|$ 使用归纳法。

基例： 如果 $|C| = 1$，则该命题显然成立。

归纳步骤： 假设对于所有大小小于 $|C|$的码，该命题均成立。设 $c_M \in C$ 是长度最大的字，长度为 $M := |c_M| = \max\{|c| : c \in C\}$。

根据归纳假设，存在一个与码 $C \setminus \{c_M\}$对应的对称树 T_0，即

$$|\mathcal{L}_k(T_0)| = \sum_{\substack{c \in C \setminus \{c_M\} \\ |c|=k}} 2^{[c]_a}.$$

令 $Ex_M(T_0)$ 表示将 T_0 的每个叶节点延伸至深度 M后得到的对称树。形式上，对于一个与根的距离为 k 的叶节点 $v \in T_0$，我们在 v处挂接一棵深度为 $M - k$ 的三叉树。等价地，就典范标记树而言，每个叶节点 $c \in T_0$ 都替换为以下新叶节点集合

$$\{c \cdot w : w \in \{a, a^{-1}, b\}^{M-|c|}\}.$$

根据构造，$Ex_M(T_0)$ 的所有叶节点到根的距离均为 M。此外，根据引理2.5，有

$$|\mathcal{L}_M(Ex_M(T_0))| = \sum_{c \in C \setminus \{c_M\}} 3^{M-|c|} \cdot 2^{[c]_a} = 3^M \sum_{c \in C \setminus \{c_M\}} 3^{-|c|} \cdot 2^{[c]_a} \leq 3^M - 2^{[c_M]_a}.$$

将深度为 M 的完全三叉树记为 $T_M(3)$，现在考虑 $T_M(3)$ 中扩展树的补：

$$Ex_M(T_0)^C := T_M(3) \setminus Ex_M(T_0).$$

Then

$$|\mathcal{L}_M(Ex_M(T_0)^C)| = 3^M - |\mathcal{L}_M(Ex_M(T_0))| \geq 2^{[c_M]_a}.$$

此外，$Ex_M(T_0)^C$ 本身就是一棵对称树。因此，它对应于满足条件的子集 $S \subseteq \{a, b\}^M$

$$\sum_{w \in S} 2^{[w]_a} \geq 2^{[c_M]_a}.$$

注意，若 $[w]_a \leq [c_M]_a$ 对所有 $w \in S$成立，则由引理 2.6，存在一个子集$S' \subseteq S$ 使得

$$\sum_{w \in S'} 2^{[w]_a} = 2^{[c_M]_a}.$$

将与 S' 对应的对称树添加到 T_0，即可得到完整码 C 所需的对称树（即具有所需数量叶节点的对称树），从而完成归纳步骤。这里，“添加”是指构造与 S' 和关联于 T_0 的前缀无关码之并集相对应的对称树。

注2.7 另请注意，一般而言，每个深度至多有一个顶点满足$[w]_a = 0$；此外，一个码至多包含一个具有此性质的字。因此，若假设 $[c_M]_a = 0$，则根据奇偶性考虑，必然存在$w \in S$ 使得 $[w]_a = 0$。在这种情况下，我们可以取$S' = \{w\}$ 作为所需的子集。

如果前一种情况不成立，则存在某个 $w \in S$ 使得 $[w]_a > [c_M]_a$。令 $i > [c_M]_a$ 为满足以下条件的最小值：存在 $w \in S$ 使得 $[w]_a = i$。

滥用记号，将 T_0 与相应的前缀无关码等同起来，我们引入记号

$$C_k(j) := \{w \in T_0 : |w| = k,\ [w]_a = j\}, \quad w_k(j) := |C_k(j)|, \qquad k \in \mathbb{N}, 0 \le j \le k$$

即， $C_k(j)$ 是 T_0 中长度为 k 且字母 a 恰好出现 j 次的字的集合。另记 $I := \{|c| : c \in T_0\}$ 为与 T_0 对应的前缀无关码中各字的长度集合。注意到

$$w_k(i) \le \binom{k}{i},$$

因为长度为 k 且恰好有 i 个字母等于 a 的序列至多有 $\binom{k}{i}$ 个。

现在，假设存在某个 $m \in I$，满足 $m < M$，使得对于某个满足下式的 j

$$\max(0, [c_M]_a - (M - m)) \le j \le \min(m, [c_M]_a),$$

我们有

$$w_m(j) < \binom{m}{j}.$$

那么就存在一个“缺失”的字

$$w \in \{a, b\}^m \setminus C_m(j), \quad \text{with } [w]_a = j.$$

我们可以通过追加.....将这个字扩展为长度为 M 的新字

$$w' := w \cdot \langle \underbrace{a, \ldots, a}_{[c_M]_a - j \text{ times}}, \underbrace{b, \ldots, b}_{M - m - ([c_M]_a - j) \text{ times}} \rangle,$$

根据构造，该字没有属于 T_0 的前缀。因此，将对应于 w' 的对称树添加到树 T_0 中，便得到完整码 C 所需的对称树，从而完成归纳步骤。

否则，令 $w \in S$ 满足 $[w]_a = i$。根据 S 的构造，必定存在某个 $m < M$ 和一个最小指标 j，使得 $\max(0, i - (M - m)) \le j \le \min(m, i)$，且

$$w_m(j) < \binom{m}{j}.$$

根据假设，令 $j' < j$ 为满足下式的最大值

$$\max(0, [c_M]_a - (M - m)) \le j' \le \min(m, [c_M]_a),$$

and

$$w_m(j') = \binom{m}{j'}, \quad w_m(j'+1) = \binom{m}{j'+1}, \ldots, w_m(j-1) = \binom{m}{j-1}.$$

在这种情况下，我们可以进行如下调整：从 $C_m(j-1)$ 中删除两个字，并将一个缺失元素添加到 $C_m(j)$ 中。由于 $w_m(j) < \binom{m}{j}$，

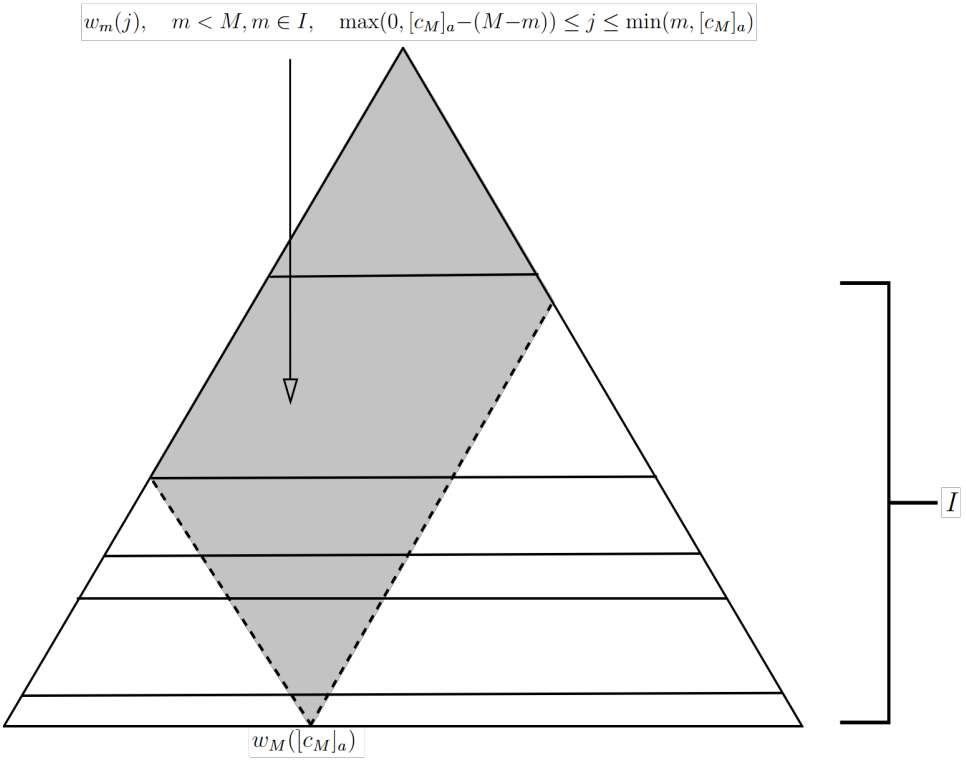

图1：我们要在其中寻找一个既不是现有码 T_0 的前缀、其自身也不以现有码 T_0 中的字为前缀的区域。

这一操作是可行的；根据构造，对称树中每种深度的叶节点总数保持不变。这可以从深度为 $m \in I$ 的叶节点总数是

$$\sum_{k=0}^{m} 2^k w_m(k).$$

我们迭代地继续这一过程。只要 $w_m(j') \ge 2$，就可以毫无问题地重复这一调整。

若 $w_m(j') = 1$，则必有 $j' = 0$。特别地，由于 j' 的极大性，这意味着 $[c_M]_a = 0$，这一点我们已在注 2.7 中讨论过。

因此，完成上述调整后，修改后的树中的叶节点总数仍与 T_0 中的相同，而 $C_m(j)$ 中的“空缺”已被“移至”这样一个区域：可以在其中添加一个包含 $[c_M]_a$ 次 a 的字。因此，可以如前述论证那样解决这一配置，从而完成所需的构造。 ■

3 进一步讨论

我们注意到，定理 1.9 可推广至任意有限字母表上的码。更确切地说，我们有如下陈述。

设 C 是一个基于大小为 A 的字母表 $\mathcal{A}$ 的有限码。那么，对于每个 $a \in \mathcal{A}$，都

存在一个前缀无关码 C' 基于 $\mathcal{A}$，使得对于每个 $k \in \mathbb{N}$ 满足 $k \geq A$[2]，

$$\sum_{\substack{w \in C \\ |w|=k}} A^{k-[w]_a} = \sum_{\substack{w' \in C' \\ |w'|=k}} A^{k-[w']_a}.$$

在这一一般情形下，我们省略证明，因为证明可由二元情形中的相同论证得出，而二元情形已经涵盖了核心思想。例如，在一般情形下，引理 2.5 指出，对于字母表 $\mathcal{A}$ 大小为 A的码，有

$$\sum_{w \in C} (2A-1)^{-|w|} A^{|w|-[w]_a} \leq 1.$$

这可通过上述相同的论证（以及定理 3 在 [2] 中证明所用的相同计数论证）得出。

此外，鉴于注记 1.10，自然会问：一个码在基数上能与满足定理 1.9结论的相应前缀无关码相差多小。更确切地说，给定一个码 C，是否总能找到一个前缀无关码 C' 满足定理 1.9 的结论，且 $\big||C|-|C'|\big| = O(|C|)$? 甚至 $\big||C|-|C'|\big| = O(1)$?

参考文献

[1] C. Cordero. 关于三角形猜想的计算机探索札记。, 语言与自动机理论及应用 (2019).[2] D. Kraizberg. 开放博弈的获胜准则：前缀码的博弈论方法。, arXiv 预印本 (2026).[3] D. Perrin, M.P. Sch¨utzenberger. 一个初等 ´的 ´信息 ´论问题。信息论 276, 法国国家科学研究中心 (1977) 249-260.[4] E. Solan. 博雷尔博弈, Routledge Taylor & Francis Group (2025).[5] J. Berstel, D. Perrin, C. Reutenauer, 码与自动机, 数学及其应用百科全书第129卷，剑桥大学出版社，剑桥 (2009).[6] M.P. Sch¨utzenberger.变长码. 手稿。Dominique Perrin，编，《第七届理论计算机科学春季学校论文集》，LITP 和 ENSTA (1980) 247–271.[7] P.W. Shor. 三角形猜想的一个反例.《组合理论杂志》A辑，38(1) (1985) 110–112.[8] R. Diestel.图论，第六版，数学研究生教材，第173卷，施普林格，柏林 (2025).[9] S. Mauceri, A. Restivo. 一族与前缀码交换等价的码。《信息处理快报》，12(1) (1981) 1–4。

[2]施加限制 $k \geq A$ 仅为确保条件 $w_k(j) < A$ 能推出 $j = 0$。