信息论 4 月 7 号第二次小测解答

中国科学技术大学《信息论 A》006125.01 班助教组

2025 年 4 月 12 日

第 1 题

Prove that a code is uniquely decodable if and only if for any integer $n \geq 1$, *and any* $\underline{s} \neq \underline{s}' \in \mathcal{S}^n$, $f(\underline{s}) \neq f(\underline{s}')$.

证明:

必要性：如果 f 是惟一可译码, n 是正整数, $\underline{s}$, $\underline{s}' \in \mathcal{S}^n$, $\underline{s} \neq \underline{s}'$, 则 $f(\underline{s}) \neq f(\underline{s}')$.

充分性：对所有正整数 n 和 $\underline{s}$, $\underline{s}' \in \mathcal{S}^n$ 有 $f(\underline{s}) \neq f(\underline{s}')$. 对所有 y_1, y_2, $\cdots$, y_m, z_1, z_2, $\cdots$, $z_n \in \mathcal{S}$, 因为

$$f(y_1)f(y_2)\cdots f(y_m)f(z_1)f(z_2)\cdots f(z_n) \neq f(z_1)f(z_2)\cdots f(z_n)f(y_1)f(y_2)\cdots f(y_m),$$

所以 $f(y_1)f(y_2)\cdots f(y_m) \neq f(z_1)f(z_2)\cdots f(z_n)$. 因此 f 是惟一可译的. $\square$

第 2 题

A code is called suffix-free if for any $s \neq s' \in \mathcal{S}, f(s)$ *is not a suffix of* $f(s')$, *and is called fix-free if it is both prefix-free and suffix-free. For a DMS* S *with* $|\mathcal{S}| < \infty$, *when* $\sum_{s\in\mathcal{S}} q^{-\ell(s)} \leq 1/2$, *find a method to construct a q-ary fix-free code with lengths* $\{\ell(s) : s \in \mathcal{S}\}$.

解:

根据题意可知，无缀码是指一个码既是无前缀的，也是无后缀的。

因为 $\sum_{s\in\mathcal{S}} q^{-\ell(s)} \leq 1/2 < 1$，即满足 Kraft 不等式，根据 Theorem 5.1 可知存在无前缀码。不妨考虑 $\ell(a_1) \leq \ell(a_2) \leq \cdots \leq \ell(a_{|\mathcal{S}|}) = \ell_{max}$，以及一个所有叶子节点深度均为 ℓ_{max} 的 q-叉树，我们先按照无前缀码的构造方式来构造：首先对于 a_1，在深度 $\ell(a_1)$ 为其分配一个节点即码字，记为 $f(a_1)$，并且删掉其所有的子节点，使其变成一个叶子节点；然后对于 a_2，不同的是，我们要在深度 $\ell(a_2)$ 的节点中找一个后缀不包含 $f(a_1)$ 的节点，记为 $f(a_2)$，

然后删掉其所有的子节点变成一个叶子节点；以此类推，直到最后一个 $a_{|\mathcal{S}|}$，我们要在深度为 $\ell(a_{|\mathcal{S}|})$ 中的节点中找一个后缀不包含 $\{f(a_1), f(a_2), \ldots, f(a_{|\mathcal{S}|})\}$ 的节点，记为 $f(a_{|\mathcal{S}|})$.

下面证明我们这种构造方式是可以构造出来一个无缀码的：

利用数学归纳法，

- 当 $\ell_{max} = 1$ 时，$\sum_{s\in\mathcal{S}} q^{-\ell(s)} = |\mathcal{S}|q^{-1} \le 1/2$，则 $|\mathcal{S}| \le 1/2 \cdot q$，故显然可按照上述构造方法构造出 $|\mathcal{S}|$ 个无缀码；
- 假设当 $\ell_{max} = 1, 2, \ldots, \ell^* - 1$ 成立，
  即考虑任意一个 $\mathcal{S}$ 满足 $\max_{s\in\mathcal{S}}\{\ell(s)\} = \ell^*$，此时有
  $$\sum_{s\in\mathcal{S},\ell(s)<\ell^*} q^{-\ell(s)} + \sum_{s\in\mathcal{S},\ell(s)=\ell^*} q^{-\ell(s)} \le 1/2 \tag{1}$$
  由于 $\sum_{s\in\mathcal{S},\ell(s)<\ell^*} q^{-\ell(s)} \le 1/2$，则存在 $s \in \mathcal{S}, \ell(s) < \ell^*$ 这些节点的无缀码。
  然后，我们考虑上述节点对 ℓ^* 层的影响，也就是说，ℓ^* 层中有多少个节点是以上述节点为前缀或者后缀的。对于前缀：我们根据讲义知道，对于长度为 $\ell(s) < \ell^*$ 的节点，在 ℓ^* 层中有 $q^{\ell^*-\ell(s)}$ 个节点是以它为前缀；同理，不难发现，对于后缀，对于长度为 $\ell(s) < \ell^*$ 的节点，在 ℓ^* 层中也有 $q^{\ell^*-\ell(s)}$ 个节点是以它为后缀的。但这两个集合可能是由交叉重复的，因此 ℓ^* 层中以上述节点为前缀或者后缀的节点数量不会超过：
  $$2\sum_{s\in\mathcal{S},\ell(s)<\ell^*} q^{\ell^*-\ell(s)} \tag{2}$$
  则 ℓ^* 层中无缀的节点数量至少为：
  $$\begin{aligned} q^{\ell^*} - 2\sum_{s\in\mathcal{S},\ell(s)<\ell^*} q^{\ell^*-\ell(s)} &= 2q^{\ell^*}(1/2 - \sum_{s\in\mathcal{S},\ell(s)<\ell^*} q^{-\ell(s)}) \\ &\ge 2q^{\ell^*} \sum_{s\in\mathcal{S},\ell(s)=\ell^*} q^{-\ell(s)} \qquad (3) \\ &= 2\sum_{s\in\mathcal{S},\ell(s)=\ell^*} \mathbf{1} \qquad (4) \end{aligned}$$
  其中 $\mathbf{1}$ 表示长度为 ℓ^* 的码字数量。又因为需要容纳的节点数量为 $\sum_{s\in\mathcal{S},\ell(s)=\ell^*} \mathbf{1}$，且同一层（长度相等的）节点不可能互为前缀或后缀，故 ℓ^* 层有足够的节点满足无缀码条件。

$\square$