

第 7 讲 信源编码

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的
Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的
Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的
Kraft 不等式

定义 1.1

关于随机变量 X 的信源编码 C 是从 X 的取值空间 $\mathcal{X}$ 到 $\mathcal{D}^*$ 的一个映射, 其中 $\mathcal{D}^*$ 表示 D 元字母表 $\mathcal{D}$ 上有限长度的字符串所构成的集合. 用 $C(x)$ 表示 x 的码字并用 $l(x)$ 表示 $C(x)$ 的长度.

定义 1.2

设随机变量 X 的概率密度函数为 $p(x)$, 定义信源编码 $C(x)$ 的期望长度 $L(C)$ 为

$$L(C) = \sum_{x \in \mathcal{X}} p(x)l(x).$$

其中 $l(x)$ 表示对应于 x 的码字长度. 并且, 对于所有 $x \in \mathcal{X}$, 令 $V_n(x) = 0$.

例 1.3

设随机变量 X 的分布及其码字分配如下：

$$P\{X = 1\} = \frac{1}{2} \quad \text{码字 } C(1) = 0$$

$$P\{X = 2\} = \frac{1}{4} \quad \text{码字 } C(2) = 10$$

$$P\{X = 3\} = \frac{1}{8} \quad \text{码字 } C(3) = 110$$

$$P\{X = 4\} = \frac{1}{8} \quad \text{码字 } C(4) = 111$$

易知 X 的熵 $H(X)$ 为 1.75 比特, 而期望长度 $L(C) = El(X)$ 也是 1.75 比特.

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

例 1.4: 设

机变量 X 的分布及其码字分配如下:

$$P\{X = 1\} = \frac{1}{3} \quad \text{码字 } C(1) = 0$$

$$P\{X = 2\} = \frac{1}{3} \quad \text{码字 } C(2) = 10$$

$$P\{X = 3\} = \frac{1}{3} \quad \text{码字 } C(3) = 11$$

易知 X 的熵 $H(X)$ 为 $\log 3 = 1.58$ 比特, 而期望长度 $L(C) = El(X)$ 是 1.67 比特, 此时 $El(x) > H(x)$.

定义 1.5

如果编码将 X 的取值空间中的每个元素映射成 $\mathcal{D}^*$ 中不同的字符串, 即

$$x \neq x' \Rightarrow C(x) \neq C(x')$$

则称这个编码是非奇异的.

定义 1.6

编码 C 的扩展 C^* 是从 $\mathcal{X}$ 上的有限长字符串到 $\mathcal{D}$ 上的有限长字符串的映射, 定义为

$$C(x_1x_2\cdots x_n) = C(x_1)C(x_2)\cdots C(x_n).$$

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

例 1.7

如果 $C(x_1) = 00$ 且 $C(x_2) = 11$, 则 $C(x_1x_2) = 0011$.

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

唯一可译码的 Kraft 不等式

定义 1.8

如果一个编码的扩展编码是非奇异的，则称该编码是惟一可译的。

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

定义 1.9

若码中无任何码字是其它码字的前缀，则称该编码为前缀码或即时码。

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

为了说明各类编码之间的不同之处，考虑如下的例子：

X	奇异的	非奇异，但非唯一可译的	唯一可译，但非即时的	即时的
1	0	0	10	0
2	0	010	00	10
3	0	01	11	110
4	0	10	110	111

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的
Kraft 不等式

下面我们介绍由 Sardinas 和 Patterson 于 1957 年设计出的一种判断唯一可译码的测试方法.

首先, 观察码 C 最短的码字是否是其它码字的前缀. 若是, 将其所有可能的尾随后缀排列出, 也就是说将其他码字序列中截去与最短码字相同的前缀部分, 余下所得的序列为尾随后缀, 将它们加入到集合 F 中. 而这些尾随后缀又可能是某些码字的前缀, 或者某些码字是这些尾随后缀的前缀, 再将由这些尾随后缀产生的新的尾随后缀列出, 加入集合 F . 然后再观察这些新的尾随后缀是否是某些码字的前缀, 或观察有否其他码字是这些新的尾随后缀的前缀, 再将新的尾随后缀列出. 依次下去直到没有一个尾随后缀是码字的前缀或没有新的尾随后缀产生为止. 这样, 首先获得的是由最短的码字能引起的所有尾随后缀. 接着, 按照上述步骤将此短的码字、…… 所有码字可能产生的尾随后缀全部列出. 由此得到所有由 C 的可能的尾随后缀组成的集合 F . 如果集合 F 中没有包含任一码字, 则可以判断此码 C 为唯一可译码.

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

例 2.1

码 $C = \{0, 10, 1100, 1110, 1011, 1101\}$ 不是唯一可译码.

有关编码的几个例子

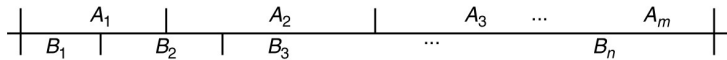
唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式



参考文献: A. A. Sardinas and C. W. Patterson, A necessary and sufficient condition for the unique decomposition of coded messages, IRE. Internat. Conv. Rec. 8 (1953), 104-108.

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的
Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

定理 3.1: Kraft 不等式

对于 D 元字母表上的即时码 (前缀码), 码字长度 $l_1, l_2, \dots, l_m$ 必定满足不等式

$$\sum_i D^{-l_i} \leq 1$$

反之, 若给定满足以上不等式的一组码字长度, 则存在一个相应的即时码, 其码字长度就是给定的长度.

证明.

考虑每一节点均含 D 个子节点的 D 叉树. 假定树枝代表码字的字符. 例如, 源于根节点的 D 条树枝代表着码字第一个字符的 D 个可能值. 另外, 每个码字均由树的一片叶子表示. 因此, 始于根节点的路径可描绘出码字中的所有字符. 码字的前缀条件表明树中无一码字是其他任一码字的祖先. 因而, 在这样的编码树中, 每一码字都去除了其可能成为码字的所有后代.

令 l_{max} 为码字集中最长码字长度. 考虑在树中 l_{max} 层的所有节点, 可知其中有些是码字, 有些是码字的后代, 而另外的节点不是码字, 也不是码字的后代. 在树中 l_i 层的码字拥有 l_{max} 层中的 $D^{l_{max}-l_i}$ 个后代. 所有这些后代互不相交. 而且, 这些集合中的总结点数必定不超过 $D^{l_{max}}$. 因此, 对所有码字求和, 可得

$$\sum D^{l_{max}-l_i} \leq D^{l_{max}}.$$

于是

$$\sum D^{-l_i} \leq 1,$$

这就是 Kraft 不等式.

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

证明.

反之, 若给定任意一组满足 Kraft 不等式的码字长度 $l_1, l_2, \dots, l_m$, 总可以构造出如下的一个编码树: 将第一个深度为 l_1 的节点 (依字典序) 标为码字 1, 同时去除其所有后代. 然后在剩余的节点中找出第一个深度为 l_2 的节点, 将其标为码字 2, 同时去除树中所有属于它的所有后代, 等等. 按此方法继续下去, 则可构造出一个码字长度为 $l_1, l_2, \dots, l_m$ 的前缀码. $\square$

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

定理 3.2: 推广的 Kraft 不等式

对任意构成前缀码的可数无限码字集, 码字长度也满足推广的 Kraft 不等式:

$$\sum_{i=1}^{\infty} D^{-l_i} \leq 1.$$

反之, 若给定任意满足推广的 Kraft 不等式的 $l_1, l_2, \dots$, 则可构造出具有相应码长的前缀码.

证明.

我们不妨设 D 元字母表为 $\{0, 1, \dots, D-1\}$, 第 i 个码字为 $y_1 y_2 \cdots y_{l_i}$. 记 $0.y_1 y_2 \cdots y_{l_i}$ 是以 D 进制表示的实值小数, 即

$$0.y_1 y_2 \cdots y_{l_i} = \sum_{j=1}^{l_i} y_j D^{-j}.$$

由此, 这个码字对应一个区间

$$[0.y_1 y_2 \cdots y_{l_i}, 0.y_1 y_2 \cdots y_{l_i} + \frac{1}{D^{l_i}}),$$

这是一个实数集合, 集合中的所有数的 D 进制展开都以 $0.y_1 y_2 \cdots y_{l_i}$ 开始. 由前缀条件我们知所有这些区间互不相交. 因此, 它们的区间长度总和小于或等于 1. 因此我们证明了

$$\sum_{i=1}^{\infty} D^{-l_i} \leq 1.$$

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

证明.

反之, 和有限情形类似。首先将长度下表重新排列, 使得 $l_1 \leq l_2 \leq \dots$. 然后从单位区间的左端开始, 依次进行分配, 即可获得满足条件的码字集. 例如, 如果想构造一个二元编码使得 $l_i = i, i = 1, 2, \dots$, 那么将区间 $[0, \frac{1}{2}), [\frac{1}{2}, \frac{3}{4}), \dots$ 分配给字符, 使其对应码字为 $0, 10, \dots$. □

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的
Kraft 不等式

前面我们已经证明了满足前缀条件的任何一个码字集合满足 Kraft 不等式, 并且当一组码字长度集合满足 Kraft 不等式时, 存在这样的码字集, 它们的长度集合正好就是给定的长度集合. 下面我们来考虑求解前缀码的最小期望长度问题.

由前面的讨论我们知该问题等价于寻找一个前缀码, 它的码字长度集合满足 Kraft 不等式, 其码字期望长度 $l = \sum p_i l_i$ 达到最小. 这就化为了一个最优化的问题: 在所有整数 $l_1, l_2, \dots, l_m$ 上, 最小化

$$L = \sum p_i l_i$$

其约束条件为

$$\sum D^{-l_i} \leq 1.$$

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

先利用微积分的知识做一个简单的分析, 以此说明达到最小值时 l^* 应具有的形式. 取消 l_i 必须是整数的限制, 并假定约束条件中的等号成立. 于是, 利用拉格朗日乘子法, 将带约束的最小化问题转化为求

$$J = \sum p_i l_i + \lambda (\sum D^{-l_i})$$

的最小化问题. 关于 l_i 求微分, 可得

$$\frac{\partial J}{\partial l_i} = p_i - \lambda D^{-l_i} \ln D.$$

令偏导数为 0, 得

$$D^{-l_i} = \frac{p_i}{\lambda \ln D}.$$

将此代入约束条件中以求得合适的 λ , 可得 $\lambda = 1/\ln D$, 因而 $p_i = D^{-l_i}$. 即最优码长为

$$l_i^* = -\log_D p_i.$$

若可以取码字长度为非整数, 则此时的期望码字长度为

$$L^* = \sum p_i l_i^* = -\sum p_i \log_D p_i = H_D(X).$$

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的
Kraft 不等式**定理 4.1**

随机变量 X 的任一 D 元即时码的期望长度必定大于等于熵 $H_D(X)$, 即

$$L \geq H_D(X)$$

当且仅当 $D^{-l_i} = p_i$, 等号成立.

证明.

我们将期望长度与熵的差写成如下形式

$$\begin{aligned} L - H_D(X) &= \sum p_i l_i - \sum p_i \log_D \frac{1}{p_i} \\ &= - \sum p_i \log_D D^{-l_i} + \sum p_i \log_D p_i \end{aligned}$$

设 $r_i = D^{-l_i} / \sum_j D^{-l_j}$, $c = \sum D^{-l_i}$, 由相对熵的非负性以及 $c \leq 1$ (利用 Kraft 不等式), 可得

$$\begin{aligned} L - H_D(X) &= \sum p_i \log_D \frac{p_i}{r_i} - \log_D c \\ &= D(\mathbf{p} \parallel \mathbf{r}) + \log_D \frac{1}{c} \\ &\geq 0. \end{aligned}$$

因此, $L \geq H$, 当且仅当 $p_i = D^{-l_i}$ (即对所有的 i , $-\log_D p_i$ 为整数), 等号成立. □

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

定义 4.2

对于某个 n , 如果概率分布的每一个概率值均等于 D^{-n} . 则称这个概率分布是 D 进制的.

因此, 当且仅当 X 的分布是 D 进制的, 上述定理等号成立.

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的
Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

定理 4.3

设 $l_1^*, l_2^*, \dots, l_m^*$ 是关于信源分布 $\mathbf{p}$ 和一个 D 元字母表的一组最优码长, L^* 为最优码的相应期望长度 ($L^* = \sum p_i l_i^*$), 则

$$H_D(X) \leq L^* < H_D(X) + 1.$$

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

证明.

设 $l_i = \lceil \log_D \frac{1}{p_i} \rceil$, 则 l_i 满足 Kraft 不等式且

$$H_D(X) \leq L = \sum p_i l_i < H_D(X) + 1.$$

但由于 L^* 是最优码的期望长度, 它不大于 $L = \sum p_i l_i$. 再由定理4.1可知 $L^* \geq H_D$, 从而得证. □

定义 L_n 为每输入字符期望码字长度, 也就是说, 如果设 $l(x_1, x_2, \dots, x_n)$ 是与 $(x_1, x_2, \dots, x_n)$ 相应的二进制码字长度 (为简便起见, 在本小节余下的部分, 假定 $D = 2$), 则

$$L_n = \frac{1}{n} p(x_1, x_2, \dots, x_n) l(x_1, x_2, \dots, x_n) = \frac{1}{n} El(X_1, X_2, \dots, X_n).$$

将前面推导的界应用于此时的编码, 有

$$H(X_1, X_2, \dots, X_n) \leq El(X_1, X_2, \dots, X_n) < H(X_1, X_2, \dots, X_n) + 1.$$

若 $X_1, X_2, \dots, X_n$ 是独立同分布的, 因此

$H(X_1, X_2, \dots, X_n) = \sum H(X_i) = nH(X)$. 上式两边同除以 n , 得

$$H(X) \leq L_n < H(X) + \frac{1}{n}.$$

因此, 通过足够大的分组长度, 可以获得一个编码, 可以使其每字符期望码长任意地接近熵.

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

若此随机过程不是独立同分布的，我们亦然可以类似讨论，我们有估计

$$H(X_1, X_2, \dots, X_n) \leq El(X_1, X_2, \dots, X_n) < H(X_1, X_2, \dots, X_n) + 1.$$

上式两边同时除以 n ，且定义 L_n 为每字符期望描述长度，可得

$$\frac{H(X_1, X_2, \dots, X_n)}{n} \leq L_n < \frac{H(X_1, X_2, \dots, X_n)}{n} + \frac{1}{n}.$$

如果随机过程是平稳的，则 $H(X_1, X_2, \dots, X_n)/n \rightarrow H(\mathcal{X})$. 于是我们有如下定理.

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

定理 4.4

每字符最小期望码字长满足

$$\frac{H(X_1, X_2, \dots, X_n)}{n} \leq L_n^* < \frac{H(X_1, X_2, \dots, X_n)}{n} + \frac{1}{n}.$$

进一步, 若 $X_1, X_2, \dots, X_n$ 是平稳随机过程, 则

$$L_n^* \rightarrow H(\mathcal{X}).$$

其中 $H(\mathcal{X})$ 为随机过程的熵率.

最后我们讨论当面对的对象非真实分布时，期望描述长度会变得怎么样？下面考虑概率密度函数 $q(x)$ 的香农编码，相应的码长为 $l(x) = \lceil \log \frac{1}{q(x)} \rceil$. 假定真实分布的概率密度函数是 $p(x)$.

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

最后我们讨论当面对的对象非真实分布时，期望描述长度会变得怎么样？下面考虑概率密度函数 $q(x)$ 的香农编码，相应的码长为 $l(x) = \lceil \log \frac{1}{q(x)} \rceil$. 假定真实分布的概率密度函数是 $p(x)$.

定理 4.6: 偏码

码字长度分配 $l(x) = \lceil \log \frac{1}{q(x)} \rceil$ 关于 $p(x)$ 的期望码长满足

$$H(p) + D(p\|q) \leq E_p l(X) < H(p) + D(p\|q) + 1.$$

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

证明.
期望码长为

$$\begin{aligned}
 El(x) &= \sum_x p(x) \lceil \log \frac{1}{q(x)} \rceil \\
 &< \sum_x p(x) (\log \frac{1}{q(x)} + 1) \\
 &= \sum_x p(x) \log \frac{p(x)}{q(x)} \frac{1}{p(x)} + 1 \\
 &= \sum_x p(x) \log \frac{p(x)}{q(x)} + \sum_x p(x) \log \frac{1}{p(x)} + 1 \\
 &= D(p\|q) + H(p) + 1.
 \end{aligned}$$

类似地，可以得到期望码长的下界.

□

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的
Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

唯一可译码的 Kraft 不等式

定理 5.1: McMillan

任意唯一可译的 D 元码的码字长度必然满足 Kraft 不等式

$$\sum D^{-l_i} \leq 1.$$

反之, 若给定满足上述不等式的一组码字长度, 则可以构造出具有同样码字长度的唯一可译码.

证明.

考虑编码 C 的 k 次扩展 C^k (即原先惟一可译码 C 的 k 次串联所形成的码). 由惟一可译的定义, 该码的 k 次扩展是非奇异的. 由于所有长度为 n 的不同 D 元串的数目仅为 D^n , 故由惟一可译性可知, 在码的 k 次扩展中, 长度为 n 的码序列数目必定不超过 D^n . 由此讨论我们来证明 Kraft 不等式. 设字符 $x \in \mathcal{X}$ 所对应的码字长度记为 $l(x)$. 对于扩展码, 码序列的长度为

$$l(x_1, x_2, \dots, x_k) = \sum_{i=1}^k l(x_i).$$

我们的目标是要证明

$$\sum_{x \in \mathcal{X}} D^{-l(x)} \leq 1.$$

有关编码的几个例子

惟一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

证明.

我们考虑上式左边的 k 次方, 从而有

$$\begin{aligned}
 \left(\sum_{x \in \mathcal{X}} D^{-l(x)} \right)^k &= \sum_{x_1 \in \mathcal{X}} \sum_{x_2 \in \mathcal{X}} \cdots \sum_{x_k \in \mathcal{X}} D^{-l(x_1)} D^{-l(x_2)} \cdots D^{-l(x_k)} \\
 &= \sum_{x_1, x_2, \dots, x_k \in \mathcal{X}^k} D^{-l(x_1)} D^{-l(x_2)} \cdots D^{-l(x_k)} \\
 &= \sum_{x^k \in \mathcal{X}^k} D^{(x^k)}.
 \end{aligned}$$

我们将上式右边按字符长度重新求和, 有

$$\sum_{x^k \in \mathcal{X}^k} D^{(x^k)} = \sum_{m=1}^{kl_{max}} a(m) D^{-m},$$

其中 l_{max} 是最大词长, $a(m)$ 为编码后码长为 m 的 x^k 的个数.

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

证明.

由码的惟一可译性, 我们有 $a(m) \leq D^m$, 从而我们有

$$\begin{aligned} \left(\sum_{x \in \mathcal{X}} D^{-l(x)} \right)^k &= \sum_{m=1}^{kl_{max}} a(m) D^{-m} \\ &\leq \sum_{m=1}^{kl_{max}} D^m D^{-m} \\ &= kl_{max}, \end{aligned}$$

从而

$$\sum_j D^{-l_j} \leq (kl_{max})^{1/k}.$$

由于上式对任意 k 成立, 当 $k \rightarrow \infty$ 时上式也成立. 由于 $(kl_{max})^{1/k} \rightarrow 1$, 我们有

$$\sum_{x \in \mathcal{X}} D^{-l(x)} \leq 1.$$

从而我们有 Kraft 不等式.

有关编码的几个例子

惟一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

惟一可译码的 Kraft 不等式

推论 5.2

可数无限字母表 $\mathcal{X}$ 上的一个唯一可译码也满足 Kraft 不等式.

有关编码的几个例子

唯一可译码的判断法

Kraft 不等式

最优码

最优码长的界

唯一可译码的 Kraft 不等式

证明.

由于唯一可译码的任一子集也是唯一可译码，因此无限码字集的有限子集亦满足 Kraft 不等式，故

$$\sum_{i=1}^{\infty} D^{-l_i} = \lim_{N \rightarrow \infty} \sum_{i=1}^N D^{-l_i} \leq 1.$$

反过来，给定满足 Kraft 的一组码字长度我们可以构造出相应的即时码，由于即时码是唯一可译的，所以我们就得到了所需要的唯一可译码。因而，McMillan 定理对于无限字母表情形同样成立。 □