香港中文大學
The Chinese University of Hong Kong

Chapter 4
Zero-Error Data Compression

© Raymond W. Yeung 2014

The Chinese University of Hong Kong

In this chapter

In this chapter

- Why $H(X)$ measures the amount of information in X?

In this chapter

- Why $H(X)$ measures the amount of information in X?
- A first look at data compression: Prefix codes

In this chapter

- Why $H(X)$ measures the amount of information in X?
- A first look at data compression: Prefix codes
- How to construct optimal prefix codes – Huffman codes?

4.1 The Entropy Bound

Definition 4.1 A D-ary source code $\mathcal{C}$ for a source random variable X is a mapping from $\mathcal{X}$ to $\mathcal{D}^*$, the set of all finite length sequences of symbols taken from a D-ary code alphabet.

Definition 4.1 A D-ary source code $\mathcal{C}$ for a source random variable X is a mapping from $\mathcal{X}$ to $\mathcal{D}^*$, the set of all finite length sequences of symbols taken from a D-ary code alphabet.

Definition 4.2 A code $\mathcal{C}$ is uniquely decodable if for any finite source sequence, the sequence of code symbols corresponding to this source sequence is different from the sequence of code symbols corresponding to any other (finite) source sequence.

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$AAD \quad \rightarrow \quad 0010$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$AAD \quad \longrightarrow \quad 0010$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$AAD \quad \rightarrow \quad 0010$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$AAD \quad \rightarrow \quad 0010$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$A\underline{A}D \quad \rightarrow \quad 0\underline{0}10$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$AA\underline{D} \quad \to \quad 0010$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$AA\underline{D} \quad \rightarrow \quad 00\underline{10}$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$\underline{AAD} \quad \rightarrow \quad 0010$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$\underline{AAD} \quad \rightarrow \quad \underline{0010}$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$\begin{aligned} AAD &\rightarrow 0010 \\ \underline{ACA} &\rightarrow 0010 \end{aligned}$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$\begin{aligned} AAD &\rightarrow 0010 \\ \underline{ACA} &\rightarrow \underline{0010} \end{aligned}$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$\begin{array}{lcl} AAD & \rightarrow & 0010 \\ ACA & \rightarrow & 0010 \\ \underline{AABA} & \rightarrow & 0010 \end{array}$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$\begin{aligned} AAD &\rightarrow 0010 \\ ACA &\rightarrow 0010 \\ \underline{AABA} &\rightarrow \underline{0010} \end{aligned}$$

Example 4.3 Let $\mathcal{X} = \{A, B, C, D\}$. Consider the code $\mathcal{C}$ defined by

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

$$\begin{aligned} AAD &\rightarrow 0010 \\ ACA &\rightarrow 0010 \\ AABA &\rightarrow 0010 \end{aligned}$$

Therefore, $\mathcal{C}$ not uniquely decodable.

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.
2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.
3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2$$
$$= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned} &\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\ &= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\ &= 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \end{aligned}$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.
2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.
3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned}
&\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\
&= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\
&= 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4}
\end{aligned}$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.
2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.
3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned} &\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\ &= \left(\underline{2^{-1}} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(\underline{2^{-1}} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\ &= \underline{4} \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \end{aligned}$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.
2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.
3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned}\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\ &= \left(\underline{2^{-1}} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + \underline{2^{-1}} + 2^{-2} + 2^{-2}\right) \\ &= \underline{4} \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4}\end{aligned}$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.
2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.
3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned}
&\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\
&= \left(2^{-1} + \underline{2^{-1}} + 2^{-2} + 2^{-2}\right) \cdot \left(\underline{2^{-1}} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\
&= \underline{4} \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4}
\end{aligned}$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.
2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.
3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned}\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 &= \left(2^{-1} + \underline{2^{-1}} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + \underline{2^{-1}} + 2^{-2} + 2^{-2}\right) \\ &= \underline{4} \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4}\end{aligned}$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned}
&\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\
&= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\
&= 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \\
&= A_2 \cdot 2^{-2} + A_3 \cdot 2^{-3} + A_4 \cdot 2^{-4},
\end{aligned}$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned} \left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\ &= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\ &= 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \\ &= A_2 \cdot 2^{-2} + A_3 \cdot 2^{-3} + A_4 \cdot 2^{-4}, \end{aligned}$$

where

$$A_2 = 4, \;\; A_3 = 8, \;\; A_4 = 4.$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned}
&\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\
&= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\
&= \underline{4} \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \\
&= \underline{A_2} \cdot 2^{-2} + A_3 \cdot 2^{-3} + A_4 \cdot 2^{-4},
\end{aligned}$$

where

$$A_2 = 4, \;\; A_3 = 8, \;\; A_4 = 4.$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned} \left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\ &= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\ &= 4 \cdot 2^{-2} + \underline{8} \cdot 2^{-3} + 4 \cdot 2^{-4} \\ &= A_2 \cdot 2^{-2} + \underline{A_3} \cdot 2^{-3} + A_4 \cdot 2^{-4}, \end{aligned}$$

where

$$A_2 = 4, \; A_3 = 8, \; A_4 = 4.$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned}\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\ &= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\ &= 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + \underline{4} \cdot 2^{-4} \\ &= A_2 \cdot 2^{-2} + A_3 \cdot 2^{-3} + \underline{A_4} \cdot 2^{-4},\end{aligned}$$

where

$$A_2 = 4, \; A_3 = 8, \; A_4 = 4.$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned} \left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 & \\ &= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\ &= 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \\ &= A_2 \cdot 2^{-2} + A_3 \cdot 2^{-3} + A_4 \cdot 2^{-4}, \end{aligned}$$

where

$$A_2 = 4, \; A_3 = 8, \; A_4 = 4.$$

5. Then $A_2 = 4$ is the total number of sequences of $N = 2$ codewords with a total length of 2 code symbols. Specifically, the 4 sequences are

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned} \left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 & \\ = \quad & \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\ = \quad & 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \\ = \quad & A_2 \cdot 2^{-2} + A_3 \cdot 2^{-3} + A_4 \cdot 2^{-4}, \end{aligned}$$

where

$$A_2 = 4, \quad A_3 = 8, \quad A_4 = 4.$$

5. Then $A_2 = 4$ is the total number of sequences of $N = 2$ codewords with a total length of 2 code symbols. Specifically, the 4 sequences are

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned} \left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\ &= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\ &= 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \\ &= A_2 \cdot 2^{-2} + A_3 \cdot 2^{-3} + A_4 \cdot 2^{-4}, \end{aligned}$$

where

$$A_2 = 4, \quad A_3 = 8, \quad A_4 = 4.$$

5. Then $A_2 = 4$ is the total number of sequences of $N = 2$ codewords with a total length of 2 code symbols. Specifically, the 4 sequences are

$$00(AA), \; 01(AB), \; 10(BA), \; 11(BB).$$

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned} \left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 & \\ = & \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\ = & 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \\ = & A_2 \cdot 2^{-2} + A_3 \cdot 2^{-3} + A_4 \cdot 2^{-4}, \end{aligned}$$

where

$$A_2 = 4, \ A_3 = 8, \ A_4 = 4.$$

5. Then $A_2 = 4$ is the total number of sequences of $N = 2$ codewords with a total length of 2 code symbols. Specifically, the 4 sequences are

$$00(AA), \ 01(AB), \ 10(BA), \ 11(BB).$$

6. Similarly, $A_3 = 8$ and $A_4 = 4$ are the total number of sequences of 2 codewords with a total length of 3 and 4 code symbols, respectively.

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned} \left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\ &= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\ &= 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \\ &= A_2 \cdot 2^{-2} + A_3 \cdot 2^{-3} + A_4 \cdot 2^{-4}, \end{aligned}$$

where

$$A_2 = 4, \quad A_3 = 8, \quad A_4 = 4.$$

5. Then $A_2 = 4$ is the total number of sequences of $N = 2$ codewords with a total length of 2 code symbols. Specifically, the 4 sequences are

$$00(AA), \ 01(AB), \ 10(BA), \ 11(BB).$$

6. Similarly, $A_3 = 8$ and $A_4 = 4$ are the total number of sequences of 2 codewords with a total length of 3 and 4 code symbols, respectively.

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned} \left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\ &= \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\ &= 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \\ &= A_2 \cdot 2^{-2} + A_3 \cdot 2^{-3} + A_4 \cdot 2^{-4}, \end{aligned}$$

where

$$A_2 = 4, \quad A_3 = 8, \quad A_4 = 4.$$

5. Then $A_2 = 4$ is the total number of sequences of $N = 2$ codewords with a total length of 2 code symbols. Specifically, the 4 sequences are

$$00(AA), \ 01(AB), \ 10(BA), \ 11(BB).$$

6. Similarly, $A_3 = 8$ and $A_4 = 4$ are the total number of sequences of 2 codewords with a total length of 3 and 4 code symbols, respectively.

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Example 4.5 In this example, we illustrate a technique used in the proof of Theorem 4.4.

1. Consider the code $\mathcal{C}$ in Example 4.3.

2. Let $l_1 = l_2 = 1$ and $l_3 = l_4 = 2$. These are the lengths of the codewords in $\mathcal{C}$.

3. Consider the polynomial

$$\sum_{k=1}^{4} 2^{-l_k} = \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right)$$

raised to the power N.

4. For $N = 2$, we have

$$\begin{aligned}
&\left(\sum_{k=1}^{4} 2^{-l_k}\right)^2 \\
&= \quad \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \cdot \left(2^{-1} + 2^{-1} + 2^{-2} + 2^{-2}\right) \\
&= \quad 4 \cdot 2^{-2} + 8 \cdot 2^{-3} + 4 \cdot 2^{-4} \\
&= \quad A_2 \cdot 2^{-2} + A_3 \cdot 2^{-3} + A_4 \cdot 2^{-4},
\end{aligned}$$

where

$$A_2 = 4, \; A_3 = 8, \; A_4 = 4.$$

5. Then $A_2 = 4$ is the total number of sequences of $N = 2$ codewords with a total length of 2 code symbols. Specifically, the 4 sequences are

$$00(AA), \; 01(AB), \; 10(BA), \; 11(BB).$$

6. Similarly, $A_3 = 8$ and $A_4 = 4$ are the total number of sequences of 2 codewords with a total length of 3 and 4 code symbols, respectively.

Exercise Verify that $A_3 = 8$ and list the 8 sequences of 2 codewords with a total length of 3 code symbols.

Example 4.3

x	$\mathcal{C}(x)$
A	0
B	1
C	01
D	10

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k} \right)^N$$

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k} \right)^N$$

$$= \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k} \right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k} \right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \tag{2}$$

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k} \right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k} \right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \tag{2}$$

where A_i is the coefficient of D^{-i} on the LHS.

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k} \right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k} \right)^N = \sum_{i=1}^{Nl_m} \underline{A_i} D^{-i}. \tag{2}$$

where A_i is the coefficient of D^{-i} on the LHS.

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \qquad (1)$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \qquad (2)$$

where A_i is the coefficient of D^{-i} on the LHS.

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \tag{2}$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \quad (1)$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k} \right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k} \right)^N = \sum_{i=1}^{N l_m} A_i D^{-i}. \quad (2)$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \tag{2}$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \tag{3}$$

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \quad (1)$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{N l_m} A_i D^{-i}. \quad (2)$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \quad (3)$$

because there are D^i distinct sequences of i code symbols.

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \quad (1)$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \quad (2)$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \quad (3)$$

because there are D^i distinct sequences of i code symbols.

5. Substituting (3) into (2), we have

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \tag{2}$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \tag{3}$$

because there are D^i distinct sequences of i code symbols.

5. Substituting (3) into (2), we have

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \tag{2}$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \tag{3}$$

because there are D^i distinct sequences of i code symbols.

5. Substituting (3) into (2), we have

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N$$

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \tag{2}$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \tag{3}$$

because there are D^i distinct sequences of i code symbols.

5. Substituting (3) into (2), we have

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N \leq \sum_{i=1}^{Nl_m} D^i D^{-i}$$

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \tag{2}$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \tag{3}$$

because there are D^i distinct sequences of i code symbols.

5. Substituting (3) into (2), we have

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N \leq \sum_{i=1}^{Nl_m} D^i D^{-i} = \sum_{i=1}^{Nl_m} 1$$

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \quad (1)$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \quad (2)$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \quad (3)$$

because there are D^i distinct sequences of i code symbols.

5. Substituting (3) into (2), we have

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N \leq \sum_{i=1}^{Nl_m} D^i D^{-i} = \sum_{i=1}^{Nl_m} 1 = Nl_m,$$

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \tag{2}$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \tag{3}$$

because there are D^i distinct sequences of i code symbols.

5. Substituting (3) into (2), we have

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N \leq \sum_{i=1}^{Nl_m} D^i D^{-i} = \sum_{i=1}^{Nl_m} 1 = Nl_m,$$

or

$$\sum_{k=1}^{m} D^{-l_k} \leq (Nl_m)^{1/N}.$$

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \tag{2}$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \tag{3}$$

because there are D^i distinct sequences of i code symbols.

5. Substituting (3) into (2), we have

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N \leq \sum_{i=1}^{Nl_m} D^i D^{-i} = \sum_{i=1}^{Nl_m} 1 = Nl_m,$$

or

$$\sum_{k=1}^{m} D^{-l_k} \leq (Nl_m)^{1/N}.$$

Since this inequality holds for any N, upon letting $N \to \infty$, we obtain (1), completing the proof.

Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \quad (1)$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}.$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{N l_m} A_i D^{-i}. \quad (2)$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \quad (3)$$

because there are D^i distinct sequences of i code symbols.

5. Substituting (3) into (2), we have

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N \leq \sum_{i=1}^{N l_m} D^i D^{-i} = \sum_{i=1}^{N l_m} 1 = N l_m,$$

or

$$\sum_{k=1}^{m} D^{-l_k} \leq (N l_m)^{1/N}.$$

Since this inequality holds for any N, upon letting $N \to \infty$, we obtain (1), completing the proof.

Expected Length

Expected Length

- Source random variable $X \sim \{p_1, p_2, \cdots, p_m\}$

Expected Length

- Source random variable $X \sim \{p_1, p_2, \cdots, p_m\}$
- Expected length of $\mathcal{C}$:

$$L = \sum_i p_i l_i$$

Expected Length

- Source random variable $X \sim \{p_1, p_2, \cdots, p_m\}$
- Expected length of $\mathcal{C}$:
$$L = \sum_i p_i l_i$$
- Intuitively, for a uniquely decodable code $\mathcal{C}$,
$$H_D(X) \leq L$$
because each D-ary symbol can carry at most 1 D-it of information.

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X).$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \quad (1)$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i \underline{l_i} = \sum_i p_i \underline{\log_D D^{l_i}}$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

2. Then

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

2. Then

$$L - H_D(X) \quad = \quad \sum_i p_i (\log_D p_i + \log_D D^{l_i})$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

2. Then

$$L - H_D(X) = \sum_i p_i (\log_D p_i + \log_D D^{l_i})$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

2. Then

$$L - H_D(X) = \sum_i p_i (\log_D p_i + \log_D D^{l_i})$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

2. Then

$$L - H_D(X) = \sum_i p_i (\log_D p_i + \log_D D^{l_i})$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

2. Then

$$L - H_D(X) = \sum_i p_i (\log_D p_i + \log_D D^{l_i})$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

2. Then

$$L - H_D(X) = \sum_i p_i (\log_D \underline{p_i} + \log_D \underline{D^{l_i}})$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

2. Then

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

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

---

2. Then

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

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i .$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

2. Then

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

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

2. Then

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

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a}$$
>
> with equality if and only if $a = 1$.

2. Then

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

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

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

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

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

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned} L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\ &= \sum_i p_i \log_D (p_i D^{l_i}) \\ &= (\ln D)^{-1} \sum_i p_i \underline{\ln(p_i D^{l_i})} \\ &\geq (\ln D)^{-1} \sum_i p_i \underline{\left(1 - \frac{1}{p_i D^{l_i}}\right)} \end{aligned} \tag{2}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2)
\end{aligned}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned} L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\ &= \sum_i p_i \log_D (p_i D^{l_i}) \\ &= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\ &\geq (\ln D)^{-1} \sum_i \cancel{p_i} \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \end{aligned}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned} L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\ &= \sum_i p_i \log_D (p_i D^{l_i}) \\ &= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\ &\geq (\ln D)^{-1} \sum_i \cancel{p_i} \left(1 - \frac{1}{\cancel{p_i} D^{l_i}}\right) \end{aligned} \tag{2}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i \cancel{p_i} \left(1 - \frac{1}{\cancel{p_i} D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i})
\end{aligned}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right]
\end{aligned}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \quad (1)$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \quad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right]
\end{aligned}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \quad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right]
\end{aligned}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \qquad (1)$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned} L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\ &= \sum_i p_i \log_D (p_i D^{l_i}) \\ &= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\ &\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\ &= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\ &= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\ &= (\ln D)^{-1} \left[1 - \underline{\sum_i D^{-l_i}}\right] \end{aligned}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right]
\end{aligned}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3)
\end{aligned}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) && (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) && (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \quad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \quad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln (p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \quad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \quad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln (p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

$$\sum_i D^{-l_i} = \sum_i D^{\log_D p_i} = \sum_i p_i = 1,$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

$$\sum_i D^{-l_i} = \sum_i D^{\log_D p_i} = \sum_i p_i = 1,$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

$$\sum_i D^{-l_i} = \sum_i D^{\log_D p_i} = \sum_i p_i = 1,$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) & (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) & (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

$$\sum_i D^{-l_i} = \sum_i D^{\log_D p_i} = \sum_i p_i = 1,$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \quad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \quad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

$$\sum_i D^{-l_i} = \sum_i D^{\log_D p_i} = \sum_i p_i = 1,$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

$$\sum_i D^{-l_i} = \sum_i D^{\log_D p_i} = \sum_i p_i = 1,$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

> **Theorem 4.4 (Kraft Inequality)**
>
> $$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

> **Corollary 2.30 (Fundamental Inequality)** For any $a > 0$,
>
> $$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$
>
> with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

$$\sum_i D^{-l_i} = \sum_i D^{\log_D p_i} = \sum_i p_i = 1,$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i(\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1}(1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

$$\sum_i D^{-l_i} = \sum_i D^{\log_D p_i} = \sum_i p_i = 1,$$

Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

Theorem 4.4 (Kraft Inequality)

$$\sum_{k=1}^{m} D^{-l_k} \leq 1.$$

Corollary 2.30 (Fundamental Inequality) For any $a > 0$,

$$\ln a \geq 1 - \frac{1}{a} \qquad (a = p_i D^{l_i})$$

with equality if and only if $a = 1$.

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

$$\sum_i D^{-l_i} = \sum_i D^{\log_D p_i} = \sum_i p_i = 1,$$

i.e., (3) is also tight. This completes the proof of the theorem.

Corollary 4.7 (Theorem 2.43) $H(X) \le \log |\mathcal{X}|$.

Proof

Corollary 4.7 (Theorem 2.43) $H(X) \leq \log |\mathcal{X}|$.

Proof

- Let $\mathcal{X} = \{0, 1, \cdots, |\mathcal{X}| - 1\}$.

Corollary 4.7 (Theorem 2.43) $H(X) \leq \log |\mathcal{X}|$.

Proof

- Let $\mathcal{X} = \{0, 1, \cdots, |\mathcal{X}| - 1\}$.
- Let $\mathcal{C}$ be the identity code, i.e.,

x	0	1	$\cdots$	$\lvert\mathcal{X}\rvert - 1$
$\mathcal{C}(x)$	0	1	$\cdots$	$\lvert\mathcal{X}\rvert - 1$

Corollary 4.7 (Theorem 2.43) $H(X) \leq \log |\mathcal{X}|$.

Proof

- Let $\mathcal{X} = \{0, 1, \cdots, |\mathcal{X}| - 1\}$.
- Let $\mathcal{C}$ be the identity code, i.e.,

x	0	1	$\cdots$	$\lvert\mathcal{X}\rvert - 1$
$\mathcal{C}(x)$	0	1	$\cdots$	$\lvert\mathcal{X}\rvert - 1$

- Evidently, $\mathcal{C}$ is an $|\mathcal{X}|$-ary uniquely decodable code, with expected length equals 1.

Corollary 4.7 (Theorem 2.43) $H(X) \leq \log |\mathcal{X}|$.

Proof

- Let $\mathcal{X} = \{0, 1, \cdots, |\mathcal{X}| - 1\}$.
- Let $\mathcal{C}$ be the identity code, i.e.,

x	0	1	$\cdots$	$\lvert\mathcal{X}\rvert - 1$
$\mathcal{C}(x)$	0	1	$\cdots$	$\lvert\mathcal{X}\rvert - 1$

- Evidently, $\mathcal{C}$ is an $|\mathcal{X}|$-ary uniquely decodable code, with expected length equals 1.
- By the entropy bound, we have

$$1 = L \geq H_{|\mathcal{X}|}(X).$$

Corollary 4.7 (Theorem 2.43) $H(X) \leq \log |\mathcal{X}|$.

Proof

- Let $\mathcal{X} = \{0, 1, \cdots, |\mathcal{X}| - 1\}$.
- Let $\mathcal{C}$ be the identity code, i.e.,

x	0	1	$\cdots$	$\lvert\mathcal{X}\rvert - 1$
$\mathcal{C}(x)$	0	1	$\cdots$	$\lvert\mathcal{X}\rvert - 1$

- Evidently, $\mathcal{C}$ is an $|\mathcal{X}|$-ary uniquely decodable code, with expected length equals 1.
- By the entropy bound, we have

$$1 = L \geq H_{|\mathcal{X}|}(X).$$

- Leaving the base unspecified, we have

$$H(X) \leq \log |\mathcal{X}|,$$

recovering Theorem 2.43.

Definition 4.8 The redundancy R of a D-ary uniquely decodable code is the difference between the expected length of the code and the entropy of the source.

By the entropy bound,

$$R = L - H_D(X) \geq 0.$$