Coding and Information Theory
Chapter 2
Optimal Codes

Xuejun Liang

The First Lecture of Chapter 2

Content of Chapter 2

2.1 Optimality

- Let S be a source and assume that the probabilities

$$p_i = \Pr(X_n = s_i) = \Pr(s_i)$$

  where

$$0 \leq p_i \leq 1, \qquad \sum_{i=1}^{q} p_i = 1.$$

- Assume code C for S has word-lengths $l_1, l_2, \ldots l_q$. Then the Average Word-Length is defined as

$$L = L(C) = \sum_{i=1}^{q} p_i l_i \,.$$

- Given r and the probability distribution (p_i), we try to find instantaneous r-ary codes C minimizing $L(C)$.
  - Such codes are called optimal or compact codes

Example 2.1

- Let S be the daily weather (as in Example 1.2)
- with $p_i = \frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ for $i = 1, 2, 3$.
- Consider two instantaneous codes
- binary code $\mathcal{C} : s_1 \mapsto 00,\ s_2 \mapsto 01,\ s_3 \mapsto 1$

  w_1 w_2 w_3
- $L(C) =$
- binary code $\mathcal{D} : s_1 \mapsto 00,\ s_2 \mapsto 1,\ s_3 \mapsto 01$

  w_1 w_2 w_3
- $L(D) =$

Example 2.1

- Let S be the daily weather (as in Example 1.2)
- with $p_i = \frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ for $i = 1, 2, 3.$
- Consider two instantaneous codes
- binary code $\mathcal{C}: s_1 \mapsto 00,\ s_2 \mapsto 01,\ s_3 \mapsto 1$

  w_1 w_2 w_3
- $L(C) =$ $1/4 * 2 + 1/2 * 2 + 1/4 * 1 = 1.75$

- binary code $\mathcal{D}: s_1 \mapsto 00,\ s_2 \mapsto 1,\ s_3 \mapsto 01$

  w_1 w_2 w_3
- $L(D) =$ $1/4 * 2 + 1/2 * 1 + 1/4 * 2 = 1.5$

Lemma and Definition

- Lemma 2.2
  - Given a source S and an integer r, the set of all average word-lengths $L(C)$ of uniquely decodable r-ary codes C for S is equal to the set of all average word-lengths $L(C)$ of instantaneous r-ary codes C for S.
  - Can be proved directly from Corollary 1.22
- Definition
  - An instantaneous r-ary code C is defined to be optimal if $L(C) = L_{min}(S)$, which is the greatest lower bound of average word-lengths.

Theorem 2.3: Each source S has an optimal r-ary code for each integer $r \geq 2$.

- Proof: There exists C $such\ that$ $L(C) = L_{min}(S)$

source-symbols: $s_1, \ldots, s_q$

Probability distribution: $p_1, \ldots, p_q$

Assume $\exists$ k $such\ that$ $p_i > 0$ for $i \leq k$, and $p_i = 0$ for $i > k$

Let p = min($p_1, \ldots, p_k$)

1. There exists an instantaneous *r-ary* code *C* for *S*

put $l_1 = \cdots = l_q = l$ for some l such that $r^l \geq q$, and apply Theorem 1.20.

2. {L(*D*) : L(*D*) ≤ L(*C*) and D is instantaneous *r-ary* code for *S* } is finite

The word-lengths $l_1, \ldots, l_k$ of D must satisfy $$l_i \leq \frac{L(\mathcal{C})}{p} \quad \text{for} \quad i = 1, \ldots, k,$$

Otherwise $$L(\mathcal{D}) = p_1 l_1 + \cdots + p_q l_q \geq p_i l_i > p\frac{L(\mathcal{C})}{p} = L(\mathcal{C}).$$

So there are only finitely many choices for the code-words $w_1, \ldots, w_k$ in *D*

Theorem 2.3: Each source S has an optimal r-ary code for each integer $r \geq 2$.

- Proof: There exists C such that $L(C) = L_{min}(S)$

source-symbols: $s_1, \ldots, s_q$

Probability distribution: $p_1, \ldots, p_q$

Assume $\exists\, k$ such that $p_i > 0$ for $i \leq k$, and $p_i = 0$ for $i > k$

Let p = min($p_1, \ldots, p_k$)

1. There exists an instantaneous *r-ary* code *C* for *S*

put $l_1 = \cdots = l_q = l$ for some l such that $r^l \geq q$, and apply Theorem 1.20.

2. {L(*D*) : L(*D*) ≤ L(*C*) and D is instantaneous *r-ary* code for *S* } is finite

The word-lengths $l_1, \ldots, l_k$ of D must satisfy $l_i \leq \frac{L(\mathcal{C})}{p}$ for $i = 1, \ldots, k$,

Otherwise $L(\mathcal{D}) = p_1 l_1 + \cdots + p_q l_q \geq p_i l_i > p \frac{L(\mathcal{C})}{p} = L(\mathcal{C})$.

So there are only finitely many choices for the code-words $w_1, \ldots, w_k$ in *D*

2.2 Binary Huffman Codes

- Let $T = Z_2 = \{0,1\}$, Given a source S, we renumber the source-symbols $s_1, \dots, s_q$, so that

$$p_1 \geq p_2 \geq \cdots \geq p_q \,.$$

- Form a reduced source S' by combining the two least-likely symbols.
- Given any binary code C' for S', we can form a binary code C for S:

$$\begin{array}{lllll}
\mathcal{S}: & s_1, & \dots & , s_{q-2}, & \underbrace{s_{q-1},\ s_q} \\
\mathcal{S}': & s_1, & \dots & , s_{q-2}, & s' \\
 & & & & \\
 & p_1, & \dots & , p_{q-2}, & \underbrace{p_{q-1},\ p_q} \\
 & p_1, & \dots & , p_{q-2}, & p' \\
 & & & & \\
\mathcal{C}: & w_1, & \dots & , w_{q-2}, & \underbrace{w'0,\ w'1} \\
\mathcal{C}': & w_1, & \dots & , w_{q-2}, & w'
\end{array}$$

Binary Huffman Codes (Cont.)

- Lemma 2.4
  - If the code C' is instantaneous then so is C.

- Huffman code for S
  - Constructed by

$$\begin{array}{l} \mathcal{S} \to \mathcal{S}' \to \cdots \to \mathcal{S}^{(q-2)} \to \mathcal{S}^{(q-1)} \\ \mathcal{C} \leftarrow \mathcal{C}' \leftarrow \cdots \leftarrow \mathcal{C}^{(q-2)} \leftarrow \mathcal{C}^{(q-1)}. \end{array}$$

  - Note: $C^{(q-1)} = \{\varepsilon\}$ and $C^{(q-2)} = \{\varepsilon 0, \varepsilon 1\} = \{0, 1\}$
  - It is instantaneous

Example 2.5

$$\mathcal{S} \to \mathcal{S}' \to \cdots \to \mathcal{S}^{(q-2)} \to \mathcal{S}^{(q-1)}$$
$$\mathcal{C} \leftarrow \mathcal{C}' \leftarrow \cdots \leftarrow \mathcal{C}^{(q-2)} \leftarrow \mathcal{C}^{(q-1)}.$$

- Let S have q = 5 symbols $s_1,\ldots,s_5$ with probabilities

  $p_i =$ 0.3, 0.2, 0.2, 0.2, 0.1.

  Compute Huffman code and $L(C)$