EE/Ma 126a Information Theory - Homework Set #5

Ling Li, ling@cs.caltech.edu

November 7, 2000

5.1 *Optimal code lengths that require one bit above entropy.* Consider a random variable $X \in \{0, 1\}$. The expected length L of the optimal code is obviously 1. The entropy of X is

$$H(X) = -p \log p - (1-p) \log(1-p),$$

where $p = p(0)$. Since

$$\lim_{p \to 0} H(X) = -\lim_{p \to 0} p \log p - 0 = 0,$$

we have for any $\epsilon > 0$, we can find some p such that $H(X) < \epsilon$. Thus

$$\boxed{H(X) - \epsilon + 1 < 1 = L}.$$

5.2 *Ternary codes that achieve the entropy bound.*

(a) Let $D = 3$ in Theorem 5.3.1 in [Cover], we get the expected length L of any instantaneous ternary code is greater than or equal to the entropy $H_3(X)$, i.e.,

$$L \geq H_3(X)$$

with equality iff $p_i = 3^{-l_i}$. That is, each symbol of X has a probability of the form 3^{-i}.$\square$

(b) Let l be the maximum length of the codewords, i.e.,

$$l = \max_i \{l_i\}.$$

Thus multiplying 3^l to both sides of

$$\sum_{i=1}^{m} 3^{-l_i} = \sum_{i=1}^{m} p_i = 1,$$

we get

$$\sum_{i=1}^{m} 3^{l-l_i} = 3^l.$$

Since 3^l and 3^{l-l_i} are odd numbers, we know that m must be odd. $\square$

5.3 *Shannon codes and Huffman codes.*

(a) Let the four symbols of X be 1, 2, 3, and 4.

length	codeword	X	Probability			
1	1	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	1
2	00	2	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	
3	010	3	$\frac{1}{4}$	$\frac{1}{3}$		
3	011	4	$\frac{1}{12}$			

This code has average length 2 bits.

(b) There's another Huffman code with average length 2 bits:

Length	Codeword	X	Probability			
2	00	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	1
2	01	2	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	
2	10	3	$\frac{1}{4}$	$\frac{1}{3}$		
2	11	4	$\frac{1}{12}$			

Thus the codeword length assignments $(1,2,3,3)$ and $(2,2,2,2)$ are both optimal.

(c) The Shannon code lengths are

$$\{\lceil \log 3 \rceil, \lceil \log 3 \rceil, \lceil \log 4 \rceil, \lceil \log 12 \rceil\} = \{2,2,2,4\}.$$

So the length for symbol 3 in Huffman code (a) exceeds the Shannon code length.

5.4 *Huffman code.*

(a) binary code.

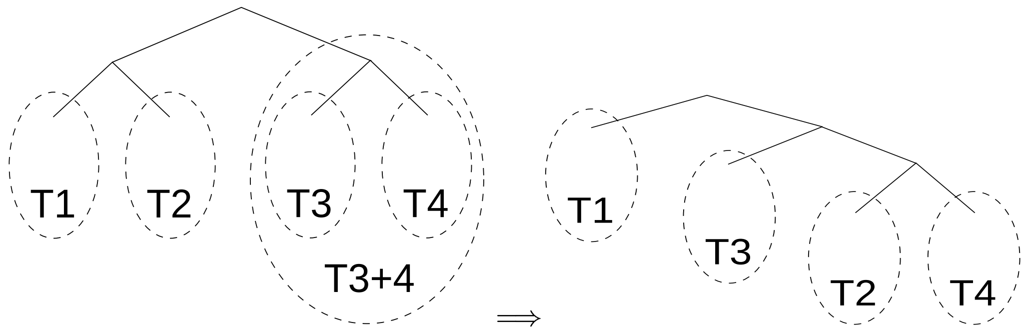
Length	Codeword	X	Probability					
2	00	1	$\frac{6}{21}$	$\frac{6}{21}$	$\frac{6}{21}$	$\frac{9}{21}$	$\frac{12}{21}$	1
2	10	2	$\frac{5}{21}$	$\frac{5}{21}$	$\frac{6}{21}$	$\frac{6}{21}$	$\frac{9}{21}$	
2	11	3	$\frac{4}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{6}{21}$		
3	010	4	$\frac{3}{21}$	$\frac{3}{21}$	$\frac{4}{21}$			
4	0110	5	$\frac{2}{21}$	$\frac{3}{21}$				
4	0111	6	$\frac{1}{21}$					

This code has average length $\boxed{\frac{17}{7} \text{ bits}}$.

(b) ternary code. Add one dummy symbol.

Length	Codeword	X	Probability			
1	1	1	$\frac{6}{21}$	$\frac{6}{21}$	$\frac{10}{21}$	1
1	2	2	$\frac{5}{21}$	$\frac{5}{21}$	$\frac{6}{21}$	
2	00	3	$\frac{4}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	
2	01	4	$\frac{3}{21}$	$\frac{3}{21}$		
3	020	5	$\frac{2}{21}$	$\frac{3}{21}$		
3	021	6	$\frac{1}{21}$			
	022	Dummy	0			

This code has an average length of $\boxed{\frac{34}{21} \text{ ternary digits}}$.



(a) Original Huffman tree (b) Transformed Huffman tree

Figure 1: Transformation of the Huffman tree.

(c) For the binary code,

$$\boxed{L = \frac{6+5+4}{21} \times 2 + \frac{3}{21} \times 3 + \frac{2+1}{21} \times 4 = \frac{51}{21} \approx 2.429}.$$

For the ternary code,

$$\boxed{L = \frac{6+5}{21} \times 1 + \frac{4+3}{21} \times 2 + \frac{2+1}{21} \times 3 = \frac{34}{21} \approx 1.619}.$$

5.5 *Optimal codeword lengths.*

(a) Assume that there is only one most probable message symbol. If not so, say, symbols α and β both are of probability $p_1 > 2/5$. Thus as long as $p_1 < 1/2$, we know that one of them must be assigned a codeword of length bigger than 1.

Suppose that the symbol with probability p_1 (denoted by α) is assigned a codeword with length bigger than 1. Thus α must appear in one of the sub-trees T_1, T_2, T_3 or T_4 in Figure 1(a). Without loss of generality, assume α is in the sub-tree T_1.

Thus from the algorithm of Huffman coding, the probability summation in sub-tree T_{3+4} (donated by $p(T_{3+4})$) should not be less than $p(T_1)$, which is not less than p_1. Since we assume that there's only one most probable symbol, the sub-tree T_{3+4} can not be only one symbol; it should contain two sub-trees T_3 and T_4. From

$$p(T_3) + p(T_4) = p(T_{3+4}) \geq p_1 > \frac{2}{5},$$

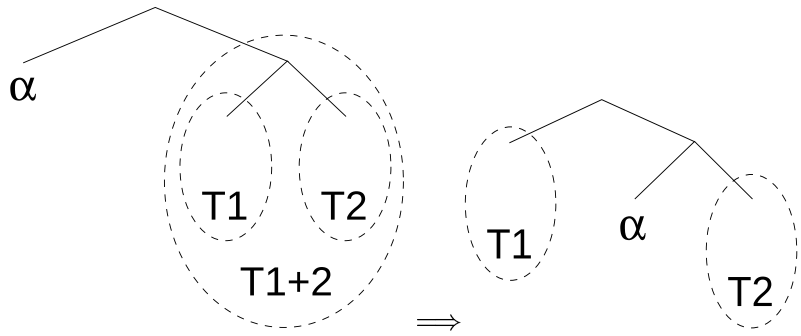
we can assume $p(T_3) > 1/5$.

Transform the Huffman tree in Figure 1(a) to the tree in Figure 1(b). The change in the expected codeword length is $p(T_2) + p(T_4) - p(T_1)$. Using $p(T_1) \geq p_1 > 2/5$, and

$$p(T_1) + p(T_2) + p(T_3) + p(T_4) = 1,$$

we get

$$p(T_2) + p(T_4) - p(T_1) = 1 - 2p(T_1) - p(T_3) < 1 - \frac{2}{5} \times 2 - \frac{1}{5} = 0,$$



(a) Original Huffman tree

(b) Transformed Huffman tree

Figure 2: Transformation of the Huffman tree.

which means that the original Huffman coding is not optimal.

Thus we know that the symbol with probability p_1 is assigned a codeword with length 1, if there is only one such symbol. $\square$

(b) Suppose that the symbol with probability $p_1 < 1/3$ (denoted by α) is assigned a codeword with length 1. Without loss of generality, assume α is the left son of the root in Figure 2(a). From $p(T_{1+2}) = 1 - p_1 > 2/3$ and p_1 is the biggest probability, we know that T_{1+2} must have two sub-trees T_1 and T_2. With

$$p(T_1) + p(T_2) = p(T_{1+2}) > \frac{2}{3},$$

we can assume $p(T_1) > 1/3$ without loss of generality.

Transform the Huffman tree in Figure 2(a) to the tree in Figure 2(b). The change in the expected codeword length is $p_1 - p(T_1) < 0$, which means that the original Huffman coding is not optimal.

Thus we know that the symbol with probability $p_1 < 1/3$ is assigned a codeword with length ≥ 2. $\square$