Solutions for chapter 2-5 in Sayood

Harald Nautsch

6th September 2004

Chapter 2

Problem 1

Since $0 \leq p_i \leq 1$, $p_i \cdot \log_2 p_i \leq 0$ which means that $H(X) = -\sum p_i \cdot \log_2 p_i \geq 0$. For the other inequality we consider $H(X) - \log_2 M$

$$\begin{aligned} H(X) - \log M &= -\sum_{i=1}^{M} p_i \log p_i - \log M \\ &= -\sum_{i=1}^{M} p_i \log p_i - \sum_{i=1}^{M} p_i \log M \\ &= \sum_{i=1}^{M} p_i \log \frac{1}{M \cdot p_i} \\ &\leq \frac{1}{\ln 2} \sum_{i=1}^{M} p_i (\frac{1}{M \cdot p_i} - 1) \\ &= \frac{1}{\ln 2} (\sum_{i=1}^{M} \frac{1}{M} - \sum_{i=1}^{M} p_i) \\ &= \frac{1}{\ln 2}(1 - 1) = 0 \end{aligned}$$

where we used the fact that $\ln x \leq x - 1$ (show this!).

Problem 3

(a) $H(X) = 2$ bits

(b) $H(X) = 1.75$ bits

(c) $H(X) \approx 1.7398$ bits

Problem 7

(a) Not uniquely decodable

(b) Not uniquely decodable

(c) Uniquely decodable

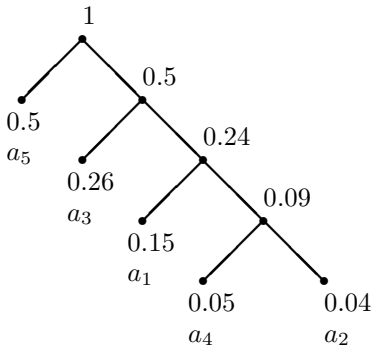
(d) Not uniquely decodable

Chapter 3

Problem 4

(a) $H = -\sum_{i=1}^{5} P(a_i) \cdot \log_2 P(a_i) \approx 1.8177$ bits

(b) The code tree will look like



The codewords can for example be:

a_1	110
a_2	1111
a_3	10
a_4	1110
a_5	0

(c) The average codeword length will be

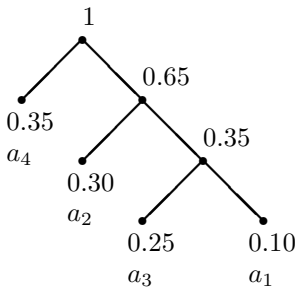
$$\bar{l} = 1 + 0.5 + 0.24 + 0.09 = 1.83 \text{ bits/codeword}$$

and the redundancy is thus

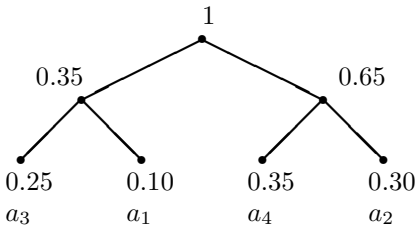
$$\bar{l} - H \approx 0.0123$$

Problem 5

(a) The code tree will look like



(b) The code tree will look like



Both codes have the same average rate (2 bits/symbol). Since the second code has codewords of the same length, it might be more useful in an environment with errors or where buffer control is needed.

Problem 13

The code will look similar to

Sequence	codeword
$a_1a_1a_1$	000
$a_1a_1a_2$	001
$a_1a_1a_3$	010
a_1a_2	011
a_1a_3	100
a_2	101
a_3	110

and have an average rate of

$$R = \frac{3}{2.19} \approx 1.3699 \text{ bits/symbol}$$

(The entropy of the source is approximately 1.1568 bits/symbol.)

Chapter 4

Problem 5

Cumulative probability function

$$F(0) = 0, \quad F(1) = 0.2, \quad F(2) = 0.5, \quad F(3) = 1$$

The first symbol is a_1

$$\begin{aligned} l^{(1)} &= 0 + (1-0) \cdot 0 = 0 \\ u^{(1)} &= 0 + (1-0) \cdot 0.2 = 0.2 \end{aligned}$$

The second symbol is a_1

$$\begin{aligned} l^{(2)} &= 0 + (0.2-0) \cdot 0 = 0 \\ u^{(2)} &= 0 + (0.2-0) \cdot 0.2 = 0.04 \end{aligned}$$

The third symbol is a_3

$$\begin{aligned} l^{(3)} &= 0 + (0.04-0) \cdot 0.5 = 0.02 \\ u^{(3)} &= 0 + (0.04-0) \cdot 1 = 0.04 \end{aligned}$$

The fourth symbol is a_2

$$\begin{aligned} l^{(4)} &= 0.02 + (0.04-0.02) \cdot 0.2 = 0.024 \\ u^{(4)} &= 0.02 + (0.04-0.02) \cdot 0.5 = 0.03 \end{aligned}$$

The fifth symbol is a_3

$$\begin{aligned} l^{(5)} &= 0.024 + (0.03-0.024) \cdot 0.5 = 0.027 \\ u^{(5)} &= 0.024 + (0.03-0.024) \cdot 1 = 0.03 \end{aligned}$$

The sixth symbol is a_1

$$\begin{aligned} l^{(6)} &= 0.027 + (0.03-0.027) \cdot 0 = 0.027 \\ u^{(6)} &= 0.027 + (0.03-0.027) \cdot 0.2 = 0.0276 \end{aligned}$$

The tag should be a number in the interval $[0.027,\ 0.0276)$, for instance we can choose the midpoint 0.0273.

Problem 6

The decoded sequence is

$$a_3a_2a_2a_1a_2a_1a_3a_2a_2a_3$$

Chapter 5

In these solutions, the symbol _ is used to denote the space character.

Problem 3

index	string	index	string	index	string	index	string
1	a	7	_b	13	ra	19	rr
2	b	8	ba	14	ay	20	ray
3	r	9	ar	15	y_	21	ya
4	y	10	r_	16	_by	22	ar_
5	_	11	_a	17	y_b	23	_ba
6	a_	12	arr	18	bar	24	

The index sequence is

$$1, 5, 2, 1, 3, 5, 9, 3, 1, 4, 7, 15, 8, 3, 13, 4, 9, 7, 14$$

Problem 4

index	string	index	string	index	string	index	string
1	a	8	hi	15	_i	22	t_is
2	_	9	is	16	is_	23	s_h
3	h	10	s_	17	_hi	24	his
4	i	11	_h	18	is_h	25	s_ha
5	s	12	ha	19	hat	26	at?
6	t	13	at	20	t_i	27	
7	th	14	t_	21	it	28	

The decoded sequence is: *this_hat_is_his_hat_it_is_his_hat*

Problem 5

index	string	index	string	index	string	index	string
1	a	6	at	11	_a	16	at_
2	_	7	ta	12	a_	17	_a_
3	r	8	ata	13	_r	18	_ra
4	t	9	atat	14	rat	19	at?
5	ra	10	t_	15	t_a	20	

The decoded sequence is: *ratatatat_a_rat_at_a_rat*

Problem 6

The resulting sequence of triples is:
$< 0, 0, 2 >$ $< 0, 0, 1 >$ $< 0, 0, 4 >$ $< 1, 1, 1 >$ $< 0, 0, 5 >$ $< 5, 2, 3 >$
$< 9, 3, 3 >$ $< 4, 1, 5 >$ $< 7, 4, 4 >$ $< 3, 1, 5 >$ $< 12, 4, 1 >$

Problem 7

The decoded sequence is: *ratatatatat_a_rat_at_a_rat*