Coding info theory, CH1 Solutions

CHAPTER 1

1.2.1

a) $|A| = 4$ b) $|A| = 3$ c) $|A| = 50$

d) The prime numbers between 2 and 40 are

$$A = \{3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37\}$$

so

$$|A| = 11$$

1.2.2

a) $A \cup B = \{0, 1, 2, 3, 4, 6, 8\} \Rightarrow |A \cup B| = 7$

b) $A \cap B = \{0, 2, 4\} \Rightarrow |A \cap B| = 3$

1.2.3

a) $A - B = \{b, c, d, 0, 1, 3\} \Rightarrow |A - B| = 6$

b) $B - A = \{+, /, \#\} \Rightarrow |B - A| = 3$

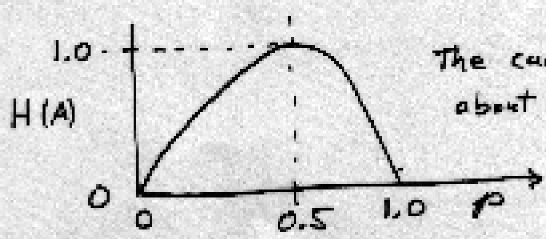
1.2.4

$$H(A) = -0.9 \log_2(0.9) - 0.1 \log_2(0.1) = 0.469 \text{ bits}$$

1.2.5

$$H(A) = p \log_2(1/p) + (1-p) \log_2\left(\frac{1}{1-p}\right)$$

$$= \log_2\left(\frac{1}{1-p}\right) + p \log_2\left(\frac{1-p}{p}\right)$$



The curve is symmetric about $p = 0.5$

Coding info theory, CH1 Solutions

1.2.6

$|A| = 8$ implies the source data can be represented by $\log_2(8) = 3$ bits. The entropy is

$$H(A) = -\sum_{i=0}^{7} p_i \log_2(p_i) = 2.7979$$

The efficiency is therefore $\frac{2.7979}{3} \times 100\% = 93.26\%$

1.2.7

a) $$Pr(a_i) = \sum_{j=0}^{2} Pr(a_i, b_j)$$

$\Rightarrow Pr(a_0) = 0.31 \quad Pr(a_1) = 0.17 \quad Pr(a_2) = 0.31$

$Pr(a_3) = 0.21$

$\therefore$

$$H(A) = -[.31 \log_2(.31) + .17 \log_2(.17) + .31 \log_2(.31) + .21 \log_2(.21)] = 1.955$$

b) $$Pr(b_j) = \sum_{i=0}^{3} Pr(a_i, b_j)$$

$\Rightarrow Pr(b_0) = 0.31 \quad Pr(b_1) = 0.27 \quad Pr(b_2) = 0.42$

$\therefore$

$H(B) = 1.5595$

c) $$H(A,B) = -\sum_{i=0}^{3} \sum_{j=0}^{2} p_{i,j} \log_2(p_{i,j}) = 3.4467$$

Coding info theory, CH1 Solutions

1.2.8

Since $H(A,B) = H(B,A)$, we have

$$H(A,B) = H(B) + H(A|B) \quad \text{from the chain rule}$$

so

$$H(A|B) = 3.4467 - 1.5595 = 1.8872$$

Likewise

$$H(A,B) = H(A) + H(B|A)$$

so

$$H(B|A) = 3.4467 - 1.9550 = 1.4917$$

1.2.9

Let $a_i \in A$ and $b_j \in B$. From equation (1.2.6),

$$H(A,B) = \sum_i \sum_j p_{i,j} \log_2 \left(\frac{1}{p_{i,j}} \right)$$

Likewise

$$H(B,A) = \sum_j \sum_i p_{j,i} \log_2 \left(\frac{1}{p_{j,i}} \right)$$

But since $Pr(a_i, b_j) = Pr(b_j, a_i)$

$$H(B,A) = \sum_i \sum_j p_{i,j} \log_2 \left(\frac{1}{p_{i,j}} \right) = H(A,B)$$

QED

1.2.10

$$H(B,A) = H(B) + H(A|B) \quad \text{by the chain rule}$$

From exercise 1.2.9, we know

$$H(A,B) = H(B,A) = H(B) + H(A|B)$$

QED

Coding info theory, CH1 Solutions

Coding info theory, CH1 Solutions

$\boxed{1.2.11}$

From example 1.2.2, we know $H(A)$ is maximized if all the symbol probabilities equal $1/|A| = 0.10$

Therefore

$$H(A) \leq \log_2 |A| = \log_2 (10) = 3.3219$$

The lower bound occurs if some $a_i \in A$ has $p_i = 1$ $\Rightarrow p_j = 0$ for $j \neq i$ $\Rightarrow H(A) = 0$. Therefore

$$0 \leq H(A) \leq 3.3219$$

$\boxed{1.2.12}$

$$H(A) \leq -\sum_i p_i \log_2(p_i)$$ with equality only if all the symbols are independent. Therefore

$$H(A) \leq 4.08 \text{ bits / letter}$$

$\boxed{1.2.13}$

Suppose we have a text consisting of n letters. From equation (1.2.12),

$$H(A_0, A_1, \ldots A_{n-1}) \leq n H(A)$$

The average number of bits per letter, i.e., the entropy rate, is therefore

$$\lim_{n \to \infty} \frac{H(A_0, A_1, \ldots, A_{n-1})}{n} \leq \frac{n H(A)}{n} = H(A)$$

Therefore, $H(A)$ is an upper bound

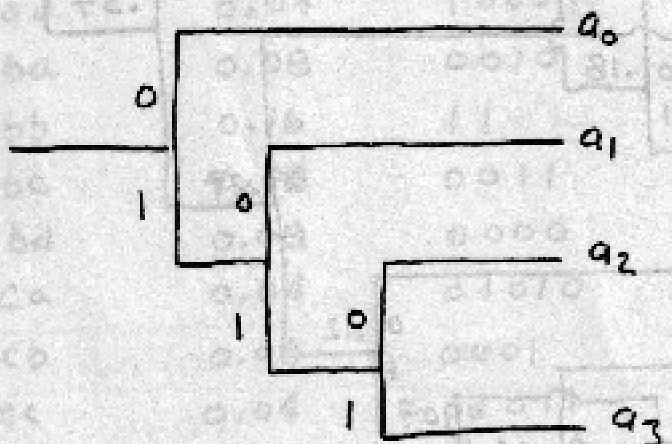
Coding info theory, CH1 Solutions

1.3.1

Using the code defined in example 1.3.1, we have the code sequence 0 0 1 0 0 1 0 0 0 1 1 0 1 1 1

1.3.2

We can decode this using a decoding tree



so

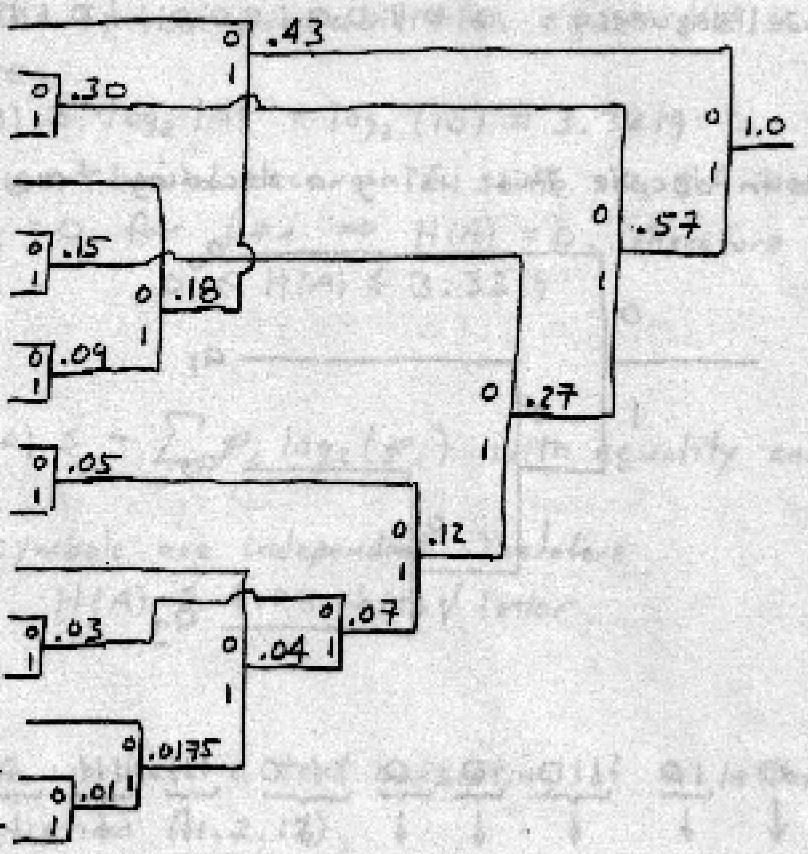
$$\underbrace{0}_{\downarrow\; a_0}\ \underbrace{10}_{\downarrow\; a_1}\ \underbrace{110}_{\downarrow\; a_2}\ \underbrace{0}_{\downarrow\; a_0}\ \underbrace{0}_{\downarrow\; a_0}\ \underbrace{110}_{\downarrow\; a_2}\ \underbrace{111}_{\downarrow\; a_3}\ \underbrace{111}_{\downarrow\; a_3}\ \underbrace{0}_{\downarrow\; a_0}$$

Coding info theory, CH1 Solutions

Coding info theory, CH1 Solutions

1.4.1

$a_0 a_0$	.25	00
$a_0 a_1$	.15	100
$a_1 a_0$	.15	101
$a_1 a_1$	.09	010
$a_0 a_2$	.075	1100
$a_2 a_0$	.075	1101
$a_1 a_2$	.045	0110
$a_2 a_1$	.045	0111
$a_0 a_3$	.025	11100
$a_3 a_0$	.025	11101
$a_2 a_2$	.0225	111110
$a_1 a_3$	.015	111100
$a_3 a_1$	.015	111101
$a_2 a_3$	.0075	1111110
$a_3 a_2$	.0075	11111110
$a_3 a_3$	.0025	11111111

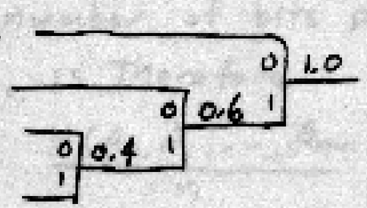


1.4.2

a) $H(A) = 0.4 \log_2\left(\frac{1}{0.4}\right) + 3(0.2)\log_2\left(\frac{1}{0.2}\right) = 1.9219$

b)

b	0.4	0
a	0.2	10
c	0.2	110
d	0.2	111



c) $\bar{L} = 0.4(1) + 0.2(2) + 0.2(3) + 0.2(3) = 2.0$

efficiency = 96.1%

Coding info theory, CH1 Solutions

1.4.3

Huffman codes are not unique. One possible code is

symbol	Prob.	Codeword	ℓ
aa	0.04	1001	0.16
ab	0.08	101	0.24
ac	0.04	10000	0.20
ad	0.04	10001	0.20
ba	0.08	0010	0.32
bb	0.16	11	0.32
bc	0.08	0011	0.32
bd	0.08	0000	0.32
ca	0.04	01010	0.20
cb	0.08	0001	0.32
cc	0.04	01011	0.20
cd	0.04	01000	0.20
da	0.04	01001	0.20
db	0.08	0110	0.32
dc	0.04	01110	0.20
dd	0.04	01111	0.20
		$\bar{L} =$	3.92

$$\text{Efficiency} = \frac{3.8438}{3.92} \times 100\% = 98.06\%$$

Your codewords may differ, but your $\bar{L}$ and your efficiency should be the same.

Coding info theory, CH1 Solutions

1.4.4

Your program should get $\bar{L} = 4.1195$

1.4.5

Let the source symbols have binary representation

a_0 : 00
a_1 : 01
a_2 : 10
a_3 : 11

The ROM table is then

Address	CONTENTS	
	Length	codeword
00	1	0xx
01	2	10x
10	3	110
11	3	111

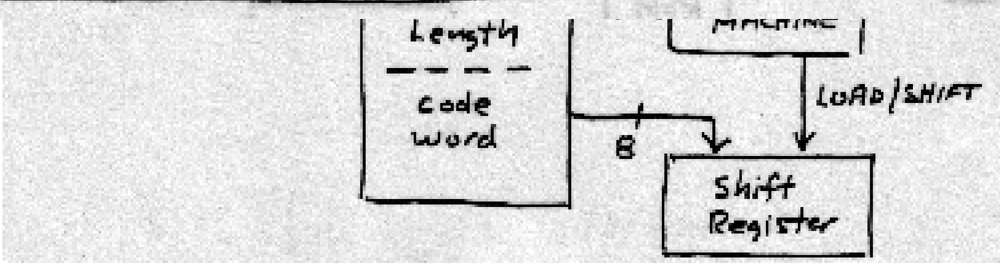
Coding info theory, CH1 Solutions

1.5.1 Following the LZ algorithm, we have the following: begin with n = 0

symbol	code	Xmit	n
b	<0,b>	—	2
c	<2,c>	2	3
c	<3,c>	3	3
a	<3,a>	3	1
c	<1,c>	1	3
b	<3,b>	3	2
c	<2,c>	—	4
c	<4,c>	4	3
c	<3,c>	—	5
c	<5,c>	5	3
c	<3,c>	—	5
c	<5,c>	—	10
c	<10,c>	10	3
c	<3,c>	—	5
c	<5,c>	—	10
c	<10,c>	—	11
c	<11,c>	11	3
a	<3,a>	—	6
c	<6,c>	6	3
c	<3,c>	—	5
c	<5,c>	—	10
a	<10,a>	10	1

DICTIONARY

Address	entry
0	<0, null>
1	<0,a>
2	<0,b>
3	<0,c>
4	<2,c>
5	<3,c>
6	<3,a>
7	<1,c>
8	<3,b>
9	<4,c>
10	<5,c>
11	<10,c>
12	<11,c>
13	<6,c>
14	<10,a>



1.5.2

rcvd symbol	Pointers	Decode
2	$\langle 0,b\rangle \to \langle 0,\text{null}\rangle$	b
3	$\langle 0,c\rangle \to \langle 0,\text{null}\rangle$	c
3	$\langle 0,c\rangle \to \langle 0,\text{null}\rangle$	c
1	$\langle 0,a\rangle \to \langle 0,\text{null}\rangle$	a
3	$\langle 0,c\rangle \to \langle 0,\text{null}\rangle$	c
4	$\langle 2,c\rangle \to \langle 0,b\rangle$	bc
5	$\langle 3,c\rangle \to \langle 0,c\rangle$	cc
10	$\langle 5,?\rangle \to \langle 3,c\rangle \to \langle 0,c\rangle$	ccc
11	$\langle 10,?\rangle \to \langle 5,c\rangle \to \langle 3,c\rangle \to \langle 0,c\rangle \Rightarrow$ cccc	
6	$\langle 3,a\rangle \to \langle 0,c\rangle \to \langle 0,\text{null}\rangle$	ca
10	$\langle 5,c\rangle \to \langle 3,c\rangle \to \langle 0,c\rangle$	ccc

DICTIONARY

address	entry
0	$\langle 0,\text{null}\rangle$
1	$\langle 0,a\rangle$
2	$\langle 0,b\rangle$
3	$\langle 0,c\rangle$
4	$\langle 2,c\rangle$
5	$\langle 3,c\rangle$
6	$\langle 3,a\rangle$
7	$\langle 1,c\rangle$
8	$\langle 3,b\rangle$
9	$\langle 4,c\rangle$
10	$\langle 5,c\rangle$
11	$\langle 10,c\rangle$
12	$\langle 11,c\rangle$
13	$\langle 6,c\rangle$
14	$\langle 10,-\rangle$

OUTPUT : b c c a c bc cc ccc cccc ca ccc