l_i^* are integers when p_i are powers of 2, i.e., they are of the form

$$p_i = 2^{-m}$$

This was proposed by Shannon, The Mathematical Theory of Communication. 1948

Shannon Code $\log \frac{1}{p_i}$

X	n_i	p_i	$\lceil -\log_2 p_i \rceil$	C
A	15	0.385	2	00
B	7	0.179	3	010
C	6	0.154	3	011
D	6	0.154	3	100
E	5	0.128	3	101
total	39	1		

lexicographic ordering

$$2^{-2} + 2^{-3} + 2^{-3} + 2^{-3} + 2^{-3} \leq 1$$

$$\frac{1}{4} + 4 \times \frac{1}{8} = \frac{6}{8} = \frac{3}{4} \leq 1$$

x	p_i	l_i	c_i	c_i in binary	C
A	0.385	2	0	0.00000	00
B	0.179	3	0.385	0.01100...	011
C	0.154	3	0.564	0.10010...	100
D	0.154	3	0.718	0.10110...	101
E	0.128	3	0.872	0.11011...	110

$$c_i = \sum_{j<i} p_j$$

$L(C) =$

$H(X) =$

$$\log_2 \frac{1}{p_i} \le l_i < \log_2 \frac{1}{p_i} + 1$$

$$p_i \log_2 \frac{1}{p_i} \le p_i l_i < p_i \log_2 \frac{1}{p_i} + p_i$$

$$\sum \qquad \sum \qquad \sum$$

$$H(X) \le L(C) < H(X) + 1$$

The average length of the code is at-most 1 bit away from the entropy

$L(C) = 2.615$ bits

$H(X) = 2.185$ bits

$L(C) = 0.385 * 2 +$

$(0.179 + 0.154 + 0.154 + 0.128)$

$* 3$

$= 0.385 * 2$

$+ (1 - 0.385) * 3$

$= 3 - 0.385$

$= 2.615$

$2.185 \leq \underline{2.615} < 3.185$

Another method

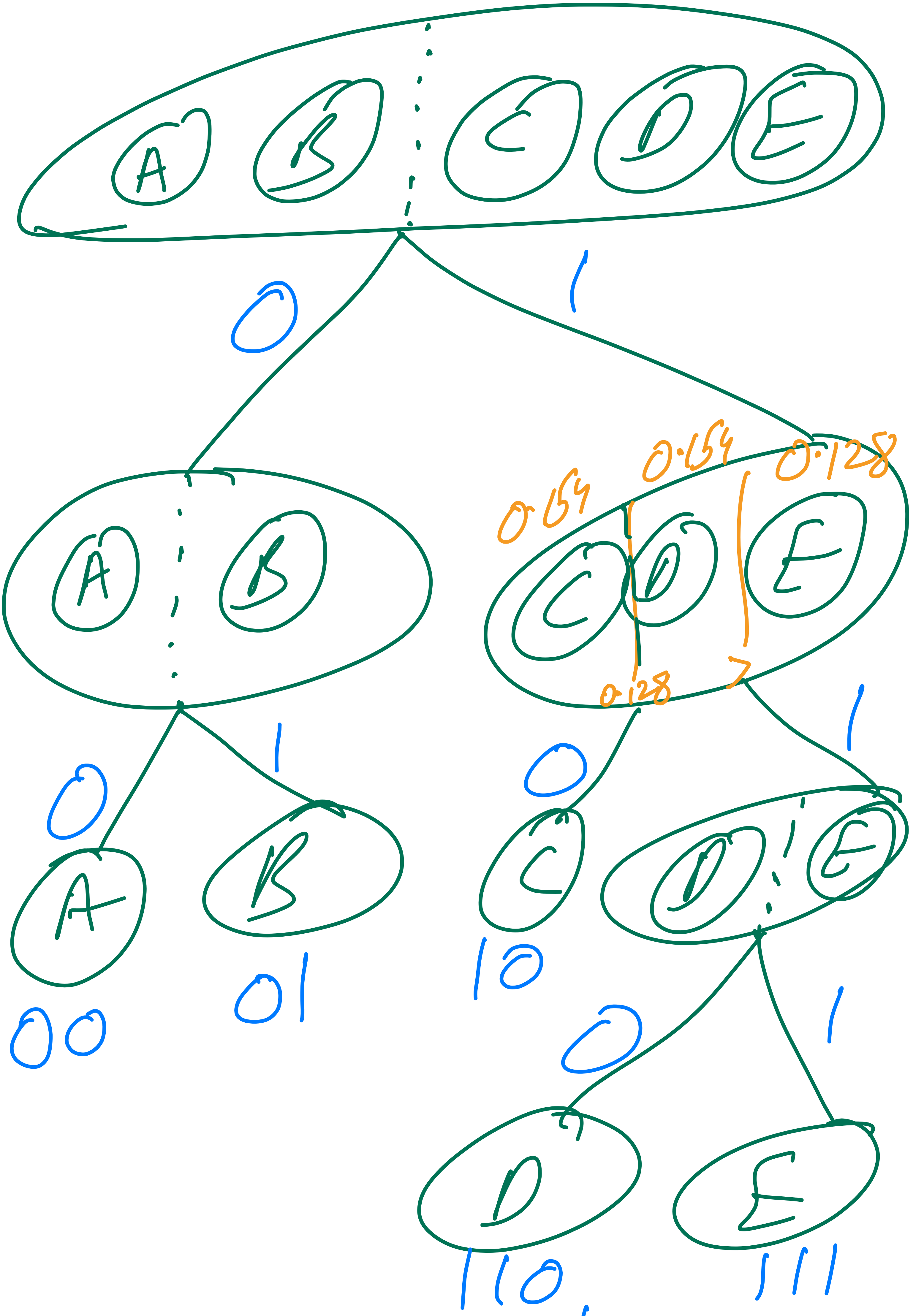
A	B	C	D	E
0.385	0.179	0.154	0.154	0.128

0.385 | 0.615

0.564 | 0.436 ✓

0.718 | 0.282

Fano's code



A B C D E
0
1
A B
C D E
0.154
0.154
0.128
0.128
7
0
1
0
1
A
B
C
D E
00
01
10
0
1
D
E
110
111

	p_i	c_i	l_i
A	0.385	00	2
B	0.179	01	2
C	0.154	10	2
D	0.154	110	3
E	0.128	111	3

$$L(C) = 0.385 * 2 + 0.179 * 2 + 0.154 * 2 + 0.154 * 3 + 0.128 * 3 =$$

2.282 bits

$2.185 < 2.282 < 2.615$

$H(X)$

Fano's code 1949

Shannon's code 1948

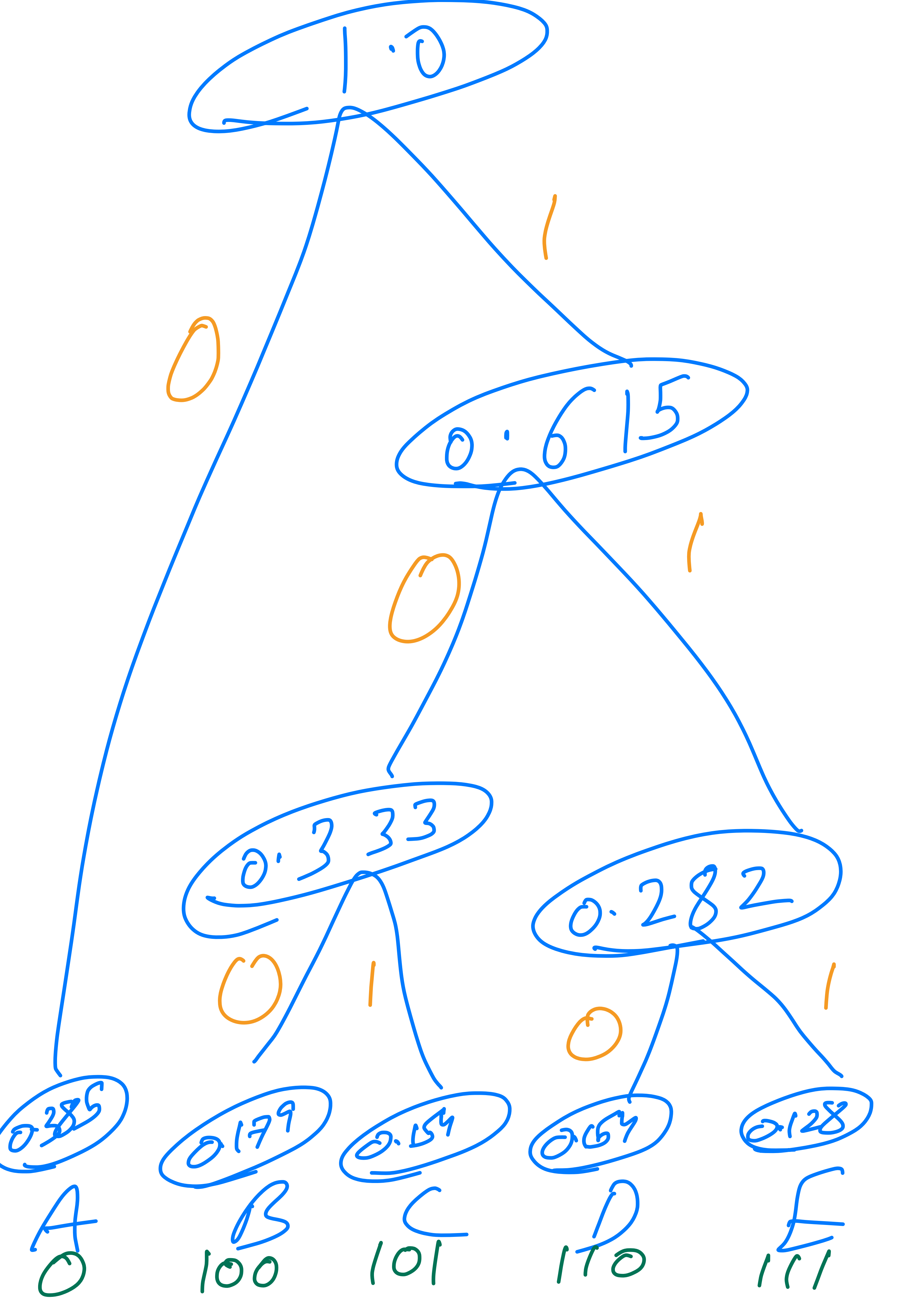
Huffman Code

$2.185 < 2.23 < 2.282 < 2.615$

$H(X)$ Huffman Fano Shannon

ZIP

Best possible



1.0
0
1
0.615
0
1
0.333
0.282
0
1
0
1
0.385
0.179
0.154
0.154
0.128
A
B
C
D
E
0
100
101
110
111

x	p_i	C	l_i
A	0.385	0	1
B	0.179	100	3
C	0.154	101	3
D	0.154	110	3
E	0.128	111	3

$$L(C) = 1 \times 0.385 + 3 \times (0.179 + 0.154 + 0.154 + 0.128)$$

$$= 0.385 + 3 * (1 - 0.385)$$

$$= 3 - 2 * 0.385$$

$$= 3 - 0.770$$

$$= 2.23 \text{ bits}$$

Shannon-Fano-Elias code

$$H(X)+1 \le L(C) < H(X)+2$$

$$F(x) = \sum_{x_i < x} p(x_i) + \frac{1}{2}p(x)$$

probabilities before x

$\frac{1}{2}$ of current

$$l_i(x) = \left\lceil \log_2 \frac{1}{p(x)} \right\rceil + 1$$

X	p_i	$\bar{F}(x)$	→ binary	l_i	
A	1/3	$\frac{1}{2}\cdot\frac{1}{3}=\frac{1}{6}$	$0.\underline{001}010$	3	001
B	1/4	$\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{4}$	$0.\underline{011}1010\ldots$	3	011
C	1/6	$\frac{1}{3}+\frac{1}{4}+\frac{1}{2}\cdot\frac{1}{6}$ $\boxed{\frac{23}{24}}$	$0.\underline{1010}10\ldots$	4	1010
D	1/4	$\frac{1}{3}+\frac{1}{4}+\frac{1}{6}+\frac{1}{2}\cdot\frac{1}{4}$ $= 0.875$	$0.\underline{111}000\ldots$	3	111

$$L(C) = 3*\frac{1}{3} + 3*\frac{1}{4} + 4*\frac{1}{6} + 3*\frac{1}{4}$$

$$= 1 + \frac{3}{2} + \frac{2}{3} = \frac{6+9+4}{6}$$

$$= \frac{19}{6} = 3.1666 \text{ bits}$$

$$H(X) = \frac{1}{3}\log 3 + \underline{\frac{2}{4}\log 4} + \frac{1}{6}\log 6$$

$$= \frac{1}{3}\log 3 + 1 + \frac{1}{6}(1 + \log 3)$$

$$= \frac{7}{6} + \frac{1}{2} * \log 3$$

$$= 1.1666 + 0.5 \times 1.585$$

$$= 1.959 \text{ bits}$$

$$\underset{H(X)+1}{2.959} \leq \underset{L(C)}{3.167} < \underset{H(X)+2}{3.959}$$