

Codificarea Huffman clasic

Sursa de informație (is) = (A, π) A - alfabet π - probabilitatea de apariție

A $a_1 \dots a_{m-1} a_m$

π $\pi_1 \dots \pi_{n-1} \pi_n$

h $\pi_1 + \dots + \pi_{m-1} + \pi_m = 1$

$\pi_1 \geq \dots \geq \pi_m$

$$L_h(is) = \sum_{i=1}^m (\pi_i \cdot |h(a_i)|) = \sum_{a \in A} (\pi(a) \cdot |h(a)|) \text{ lungimea medie a lui } h$$

Ex anaaremere

is = (A, π)

codificarea H.b.

* ~~00 110 00 00 10 01 111 01 10 01~~
a n a a r e ...

is: *

A	a	e	r	n	m
π	$\frac{3}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

0
1

 *

→ ord. alf. * → ord. apariției caracterelor în text

0 → 1; → 0 m. de bit ai codificării caracterului
frecvența caracterului

$$L_h(is) = 2 \cdot \frac{3}{10} + 2 \cdot \frac{3}{10} + 2 \cdot \frac{2}{10} + 3 \cdot \frac{1}{10} + 3 \cdot \frac{1}{10}$$

a e r n m

00 a 3 — 1 4 — 0 6

01 e 3 — 3 10 — 3 01 — 1 4 — 10 v

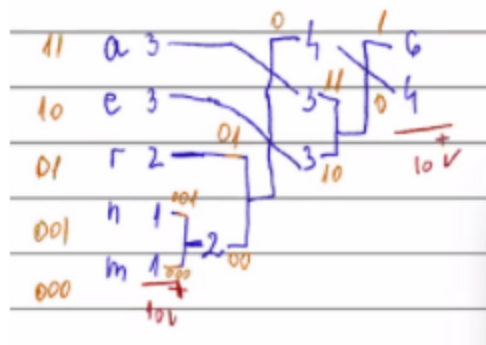
10 r 2 — 10 — 3 01 — 1 4 — 10 v

110 n 1 — 110 — 2 11 — 1 4 — 10 v

111 m 1 — 111 — 2 11 — 1 4 — 10 v

10 v
total de caractere

$$L_h(is) = \frac{22}{10} = 2,2$$



Ex1 Codificați folosind c.H. clasic textul
„unudeimailamare”

- extrageți sursa de informație
- obțineți o codificare Huffman binară
- calc. lungimea medie pentru hamonorfismul obținut la punctul(b).

Ex2 Decodificați 01001101010110001100110001001000011110111100111001000

is:	A	a	b	d	e	i	m	o	r	u	v
π		2	1	4	1	3	2	1	1	1	1

/17

0
1

ordinea apariției în is
(pt. caracterele care au aceeași frecv.)

Ex3 Specificați care dintre codurile este cHb (codificare H. binară)?

is:	A	a	b	c	d	e	f	g
π		0,05	0,5	0,05	0,2	0,1	0,03	0,07
h_1		0100	1	0110	00	0111	01100	0101
h_2		1101	0	110	100	101	1111	1100