

## Példa forrás kódolásra

### A HUFFMAN KÓDOLÁS

Eltérő valószínűségű szimbólumok készletének bináris, közel optimális kódolása

1. Vonjuk össze a két legkisebb valószínűségű szimbólumot!
2. Ismételjük, amíg csak két szimbólum marad!
3. Osszuk ki a biteket visszafelé!

<b>k</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>m<sub>k</sub></b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>
<b>p<sub>k</sub></b>	<b>0,08</b>	<b>0,05</b>	<b>0,12</b>	<b>0,4</b>	<b>0,1</b>	<b>0,25</b>

1. lépés

<b>m</b>	<b>AB</b>					
<b>k</b>	<b>0,13</b>	<b>0,12</b>	<b>0,4</b>	<b>0,1</b>	<b>0,25</b>	

2. lépés

<b>m</b>	<b>AB</b>	<b>CE</b>				
<b>k</b>	<b>0,13</b>	<b>0,22</b>	<b>0,4</b>			<b>0,25</b>

3. lépés

<b>m</b>	<b>ABCE</b>					
<b>k</b>	<b>0,35</b>	<b>0,4</b>				<b>0,25</b>

4. lépés

<b>m</b>	<b>ABCEF</b>	<b>D</b>				
<b>k</b>	<b>0,6</b>	<b>0,4</b>				
<b>kód</b>	<b>0010</b>	<b>0011</b>	<b>0000</b>	<b>0001</b>	<b>01</b>	<b>01</b>

<b>m<sub>k</sub></b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>
<b>p<sub>k</sub></b>	<b>0,08</b>	<b>0,05</b>	<b>0,12</b>	<b>0,4</b>	<b>0,1</b>	<b>0,25</b>
<b>kód</b>	<b>1111</b>	<b>1110</b>	<b>1101</b>	<b>0</b>	<b>1100</b>	<b>10</b>
<b>a<sub>k</sub></b>	<b>4</b>	<b>4</b>	<b>4</b>	<b>1</b>	<b>4</b>	<b>2</b>
<b>p<sub>k</sub> · a<sub>k</sub></b>	<b>0,32</b>	<b>0,20</b>	<b>0,48</b>	<b>0,4</b>	<b>0,4</b>	<b>0,5</b>

$\sum p_k \cdot a_k = 2,30bit / szimb$  átlagos kódszóhossz (egyenletes k.: 3 bit/szimb.),

Az entrópia:

$\frac{1}{p_k}$	<b>12,5</b>	<b>20</b>	<b>8,33</b>	<b>2,5</b>	<b>10</b>	<b>4</b>
$ld \frac{1}{p_k}$	<b>3,644</b>	<b>4,322</b>	<b>3,06</b>	<b>1,322</b>	<b>3,322</b>	<b>2</b>
$p_k \cdot ld \frac{1}{p_k}$	<b>0,2915</b>	<b>0,2161</b>	<b>0,3672</b>	<b>0,5289</b>	<b>0,3322</b>	<b>0,5</b>

$$H = \sum_{k=1}^5 p_k \cdot ld \frac{1}{p_k} = 2,2359bit / szimb$$

- az átlagos kódszóhossz (felülről) jól megközelíti az entrópia (az elvi határ) értékét.

Optimális a Huffman kódolás, ha minden szimbólum előfordulási valószínűsége kettőnek (negatív) egész kitevőjű hatványa.

Egyenletesen kell kódolni, ha minden valószínűség azonos. Ilyenkor a Huffman kód megegyezik az egyenletes kóddal. (k – sorszám, m – szimbólum, p – előfordulási valószínűség, a – kódszóhossz vagy bitszám, H – információelméleti entrópia)