

## 7. Entrópia forráskódolás és adattömörítés

Kódolástechnika

## Forráskódolás és adattömörítés

Egy tipikus szövegben a különböző karakterek gyakorisága eltérő. A normál (tömörítés nélküli) kódolás során minden karaktert adott számú bittel kódolunk.

## Forráskódolás és adattömörítés

Egy tipikus szövegben a különböző karakterek gyakorisága eltérő. A normál (tömörítés nélküli) kódolás során minden karaktert adott számú bittel kódolunk.

Ha megengedünk változó hosszúságú bitsorozatokat is a kódszavaknak (a gyakoribb karakterekhez rövidebb, a ritkább karakterekhez hosszabb bitsorozatot rendelve), akkor el lehet érni alacsonyabb átlagos kódszóhosszat is.

## Forráskódolás és adattömörítés

Egy tipikus szövegben a különböző karakterek gyakorisága eltérő. A normál (tömörítés nélküli) kódolás során minden karaktert adott számú bittel kódolunk.

Ha megengedünk változó hosszúságú bitsorozatokat is a kódszavaknak (a gyakoribb karakterekhez rövidebb, a ritkább karakterekhez hosszabb bitsorozatot rendelve), akkor el lehet érni alacsonyabb átlagos kódszóhosszat is.

Feltesszük, hogy az egyes karakterek gyakorisága a szövegben ismert:

$$p_1, \dots, p_K,$$

ahol  $K$  az ábécé mérete.

## Forráskódolás és adattömörítés

Ha a  $k$ -adik karakterhez rendelt kódszó hossza  $\ell_k$ , akkor az *átlagos kódszóhossz*

$$L = \sum_{k=1}^K p_k \ell_k.$$

## Forráskódolás és adattömörítés

Ha a  $k$ -adik karakterhez rendelt kódszó hossza  $\ell_k$ , akkor az *átlagos kódszóhossz*

$$L = \sum_{k=1}^K p_k \ell_k.$$

Egy kódolás *prefix-mentes*, ha semelyik kódszó eleje nem egyezik meg egy másik kódszóval. Ez a tulajdonság a dekódolhatóság miatt fontos.

## Forráskódolás és adattömörítés

Ha a  $k$ -adik karakterhez rendelt kódszó hossza  $\ell_k$ , akkor az *átlagos kódszóhossz*

$$L = \sum_{k=1}^K p_k \ell_k.$$

Egy kódolás *prefix-mentes*, ha semelyik kódszó eleje nem egyezik meg egy másik kódszóval. Ez a tulajdonság a dekódolhatóság miatt fontos.

A forrás *entrópiája*

$$H(X) = \sum_{k=1}^K p_k \log_2(1/p_k).$$

Elméleti alsó korlát: tetszőleges prefix-mentes kódra

$$L \geq H(X),$$

és a  $H(X)/L$  arány a kód *hatékonysága*.

## Shannon–Fano kódolás

Shannon–Fano kódolás esetén a kódszóhosszak

$$l_k = \lceil \log_2(1/p_k) \rceil.$$

Készítünk egy bináris fát, amelyre a levelek mélysége  $l_1, \dots, l_K$ , a kódszavakat pedig a gyökértől a levelekhez vezető út alapján adjuk meg.



## Shannon–Fano kódolás

Shannon–Fano kódolás esetén a kódszóhosszak

$$l_k = \lceil \log_2(1/p_k) \rceil.$$

Készítünk egy bináris fát, amelyre a levelek mélysége  $l_1, \dots, l_K$ , a kódszavakat pedig a gyökértől a levelekhez vezető út alapján adjuk meg.

Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .

## Shannon–Fano kódolás

Shannon–Fano kódolás esetén a kódszóhosszak

$$\ell_k = \lceil \log_2(1/p_k) \rceil.$$

Készítünk egy bináris fát, amelyre a levelek mélysége  $\ell_1, \dots, \ell_K$ , a kódszavakat pedig a gyökértől a levelekhez vezető út alapján adjuk meg.

Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .

$$\ell_1 = \lceil \log_2(1/0.37) \rceil = 2, \quad \ell_2 = \lceil \log_2(1/0.27) \rceil = 2,$$

$$\ell_3 = \lceil \log_2(1/0.24) \rceil = 3, \quad \ell_4 = \lceil \log_2(1/0.12) \rceil = 4.$$

## Shannon–Fano kódolás

Shannon–Fano kódolás esetén a kódszóhosszak

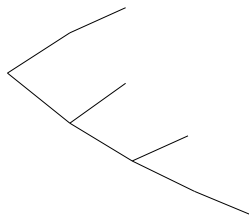
$$l_k = \lceil \log_2(1/p_k) \rceil.$$

Készítünk egy bináris fát, amelyre a levelek mélysége  $l_1, \dots, l_K$ , a kódszavakat pedig a gyökértől a levelekhez vezető út alapján adjuk meg.

Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .

$$l_1 = \lceil \log_2(1/0.37) \rceil = 2, \quad l_2 = \lceil \log_2(1/0.27) \rceil = 2,$$

$$l_3 = \lceil \log_2(1/0.24) \rceil = 3, \quad l_4 = \lceil \log_2(1/0.12) \rceil = 4.$$





## Shannon–Fano kódolás

Shannon–Fano kódolás esetén a kódszóhosszak

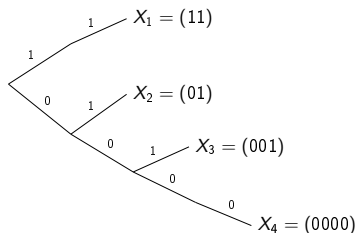
$$\ell_k = \lceil \log_2(1/p_k) \rceil.$$

Készítünk egy bináris fát, amelyre a levelek mélysége  $\ell_1, \dots, \ell_K$ , a kódszavakat pedig a gyökértől a levelekhez vezető út alapján adjuk meg.

Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .

$$\ell_1 = \lceil \log_2(1/0.37) \rceil = 2, \quad \ell_2 = \lceil \log_2(1/0.27) \rceil = 2,$$

$$\ell_3 = \lceil \log_2(1/0.24) \rceil = 3, \quad \ell_4 = \lceil \log_2(1/0.12) \rceil = 4.$$



# Shannon–Fano kódolás

Shannon–Fano kódolás esetén a kódszóhosszak

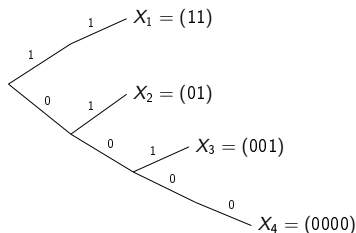
$$l_k = \lceil \log_2(1/p_k) \rceil.$$

Készítünk egy bináris fát, amelyre a levelek mélysége  $l_1, \dots, l_K$ , a kódszavakat pedig a gyökértől a levelekhez vezető út alapján adjuk meg.

Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .

$$l_1 = \lceil \log_2(1/0.37) \rceil = 2, \quad l_2 = \lceil \log_2(1/0.27) \rceil = 2,$$

$$l_3 = \lceil \log_2(1/0.24) \rceil = 3, \quad l_4 = \lceil \log_2(1/0.12) \rceil = 4.$$



Karakter	Kódszó
$X_1$	11
$X_2$	01
$X_3$	001
$X_4$	0000

# 1. feladat

Kódoljuk a következő forrást Shannon–Fano kódolással.

$$p_1 = 0.49, \quad p_2 = 0.14, \quad p_3 = 0.14, \quad p_4 = 0.07, \quad p_5 = 0.07, \\ p_6 = 0.04, \quad p_7 = 0.02, \quad p_8 = 0.02, \quad p_9 = 0.01$$

# 1. feladat

Kódoljuk a következő forrást Shannon–Fano kódolással.

$$p_1 = 0.49, \quad p_2 = 0.14, \quad p_3 = 0.14, \quad p_4 = 0.07, \quad p_5 = 0.07, \\ p_6 = 0.04, \quad p_7 = 0.02, \quad p_8 = 0.02, \quad p_9 = 0.01$$

Megoldás. A kódszó hosszak  $l_i = \lceil \log_2 1/p_i \rceil$ , tehát

$$l_1 = \lceil \log_2 1/p_1 \rceil = \lceil 1.029 \rceil = 2,$$

$$l_2 = \lceil \log_2 1/p_2 \rceil = \lceil 2.836 \rceil = 3,$$

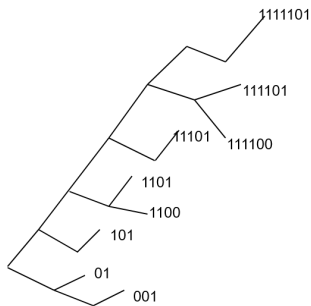
$$l_3 = \lceil \log_2 1/p_3 \rceil = \lceil 2.836 \rceil = 3,$$

$$l_4 = l_5 = 4, \quad l_6 = 5, \quad l_7 = l_8 = 6, \quad l_9 = 7.$$

Készítünk egy bináris fát, aminek 9 levele van, és a levelek mélysége éppen ezek az  $l_i$  értékek.



# 1. feladat



Szimbólum	Kódszó
$X_1$	01
$X_2$	001
$X_3$	101
$X_4$	1100
$X_5$	1101
$X_6$	11101
$X_7$	111100
$X_8$	111101
$X_9$	1111101

Megjegyzés: a kód prefix-mentes  $\Leftrightarrow$  nincs kódszó semelyik belső csúcson.

## 2. feladat

Elemezzük az előbbi kód teljesítményét.

## 2. feladat

Elemezzük az előbbi kód teljesítményét.

Megoldás. A forrás eloszlás entrópiája

$$H(X) = \sum_{i=1}^9 p_i \log_2 \left( \frac{1}{p_i} \right) = 2.314,$$

míg a kódolás átlagos kódszó hossza

$$L = 0.49 \cdot 2 + 0.28 \cdot 3 + 0.28 \cdot 3 + 0.14 \cdot 4 + \\ 0.04 \cdot 5 + 0.04 \cdot 6 + 0.01 \cdot 7 = 2.89,$$

így a kód hatékonysága

$$\frac{H(X)}{L} \approx 0.8.$$

# Huffman kódolás

A Huffman kódolás úgy építi a fát, hogy minden lépésben a két legkisebb  $p_k$  valószínűséget összevonja. A fa alapján a kódolás ugyanúgy történik, mint a Shannon–Fano kódnál.

# Huffman kódolás

A Huffman kódolás úgy építi a fát, hogy minden lépésben a két legkisebb  $p_k$  valószínűséget összevonja. A fa alapján a kódolás ugyanúgy történik, mint a Shannon–Fano kódnál.

Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .

$$p_1 = 0.37$$

$$p_2 = 0.27$$

$$p_3 = 0.24$$

$$p_4 = 0.12$$

# Huffman kódolás

A Huffman kódolás úgy építi a fát, hogy minden lépésben a két legkisebb  $p_k$  valószínűséget összevonja. A fa alapján a kódolás ugyanúgy történik, mint a Shannon–Fano kódnál.

Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .

$$p_1 = 0.37 \quad 0.37$$

$$p_2 = 0.27 \quad 0.27$$

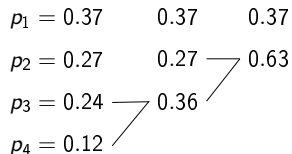
$$p_3 = 0.24 \quad \swarrow \quad 0.36$$

$$p_4 = 0.12 \quad \searrow$$

# Huffman kódolás

A Huffman kódolás úgy építi a fát, hogy minden lépésben a két legkisebb  $p_k$  valószínűséget összevonja. A fa alapján a kódolás ugyanúgy történik, mint a Shannon–Fano kódnál.

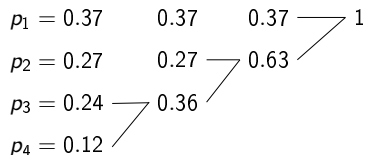
Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .



# Huffman kódolás

A Huffman kódolás úgy építi a fát, hogy minden lépésben a két legkisebb  $p_k$  valószínűséget összevonja. A fa alapján a kódolás ugyanúgy történik, mint a Shannon–Fano kódnál.

Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .

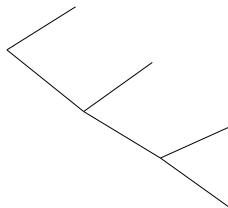
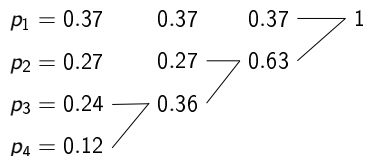




# Huffman kódolás

A Huffman kódolás úgy építi a fát, hogy minden lépésben a két legkisebb  $p_k$  valószínűséget összevonja. A fa alapján a kódolás ugyanúgy történik, mint a Shannon–Fano kódnál.

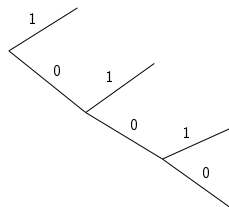
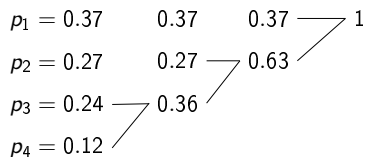
Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .



# Huffman kódolás

A Huffman kódolás úgy építi a fát, hogy minden lépésben a két legkisebb  $p_k$  valószínűséget összevonja. A fa alapján a kódolás ugyanúgy történik, mint a Shannon–Fano kódnál.

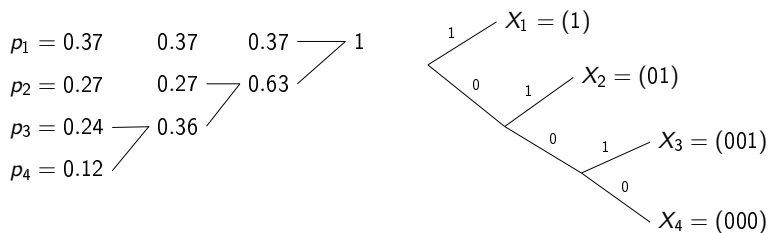
Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .



# Huffman kódolás

A Huffman kódolás úgy építi a fát, hogy minden lépésben a két legkisebb  $p_k$  valószínűséget összevonja. A fa alapján a kódolás ugyanúgy történik, mint a Shannon–Fano kódnál.

Példa.  $p_1 = 0.37$ ,  $p_2 = 0.27$ ,  $p_3 = 0.24$ ,  $p_4 = 0.12$ .



### 3. feladat

Kódoljuk a 2. feladat forrását Huffman kódolással.

### 3. feladat

Kódoljuk a 2. feladat forrását Huffman kódolással.

Megoldás. Először az állapotgráfot adjuk meg.

$$p_1 = 0.49$$

$$p_2 = 0.14$$

$$p_3 = 0.14$$

$$p_4 = 0.07$$

$$p_5 = 0.07$$

$$p_6 = 0.04$$

$$p_7 = 0.02$$

$$p_8 = 0.02$$

$$p_9 = 0.01$$

### 3. feladat

Kódoljuk a 2. feladat forrását Huffman kódolással.

Megoldás. Először az állapotgráfot adjuk meg.

$$p_1 = 0.49 \quad 0.49$$

$$p_2 = 0.14 \quad 0.14$$

$$p_3 = 0.14 \quad 0.14$$

$$p_4 = 0.07 \quad 0.07$$

$$p_5 = 0.07 \quad 0.07$$

$$p_6 = 0.04 \quad 0.04$$

$$p_7 = 0.02 \quad 0.02$$

$$p_8 = 0.02 \quad 0.03$$

$$p_9 = 0.01$$

### 3. feladat

Kódoljuk a 2. feladat forrását Huffman kódolással.

Megoldás. Először az állapotgráfot adjuk meg.

$p_1 = 0.49$	0.49	0.49
$p_2 = 0.14$	0.14	0.14
$p_3 = 0.14$	0.14	0.14
$p_4 = 0.07$	0.07	0.07
$p_5 = 0.07$	0.07	0.07
$p_6 = 0.04$	0.04	0.04
$p_7 = 0.02$	0.02	0.05
$p_8 = 0.02$	0.03	
$p_9 = 0.01$		

### 3. feladat

Kódoljuk a 2. feladat forrását Huffman kódolással.

Megoldás. Először az állapotgráfot adjuk meg.

$p_1 = 0.49$	0.49	0.49	0.49
$p_2 = 0.14$	0.14	0.14	0.14
$p_3 = 0.14$	0.14	0.14	0.14
$p_4 = 0.07$	0.07	0.07	0.07
$p_5 = 0.07$	0.07	0.07	0.07
$p_6 = 0.04$	0.04	0.04	0.09
$p_7 = 0.02$	0.02	0.05	
$p_8 = 0.02$	0.03		
$p_9 = 0.01$			



### 3. feladat

Kódoljuk a 2. feladat forrását Huffman kódolással.

Megoldás. Először az állapotgráfot adjuk meg.

$p_1 = 0.49$	0.49	0.49	0.49	0.49
$p_2 = 0.14$	0.14	0.14	0.14	0.14
$p_3 = 0.14$	0.14	0.14	0.14	0.14
$p_4 = 0.07$	0.07	0.07	0.07	0.14
$p_5 = 0.07$	0.07	0.07	0.07	0.09
$p_6 = 0.04$	0.04	0.04	0.09	
$p_7 = 0.02$	0.02	0.05		
$p_8 = 0.02$	0.03			
$p_9 = 0.01$				

### 3. feladat

Kódoljuk a 2. feladat forrását Huffman kódolással.

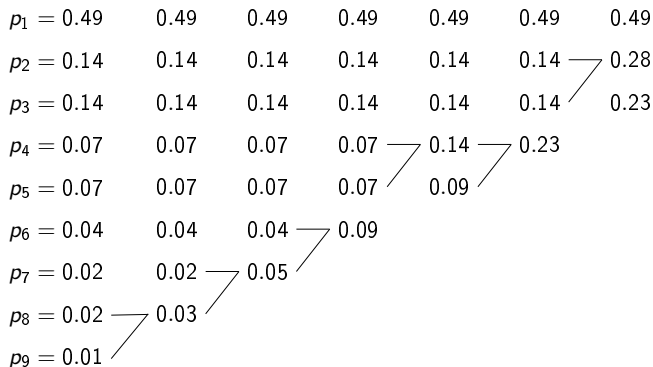
Megoldás. Először az állapotgráfot adjuk meg.

$p_1 = 0.49$	0.49	0.49	0.49	0.49	0.49
$p_2 = 0.14$	0.14	0.14	0.14	0.14	0.14
$p_3 = 0.14$	0.14	0.14	0.14	0.14	0.14
$p_4 = 0.07$	0.07	0.07	0.07	0.14	0.23
$p_5 = 0.07$	0.07	0.07	0.07	0.09	
$p_6 = 0.04$	0.04	0.04	0.09		
$p_7 = 0.02$	0.02	0.05			
$p_8 = 0.02$	0.03				
$p_9 = 0.01$					

### 3. feladat

Kódoljuk a 2. feladat forrását Huffman kódolással.

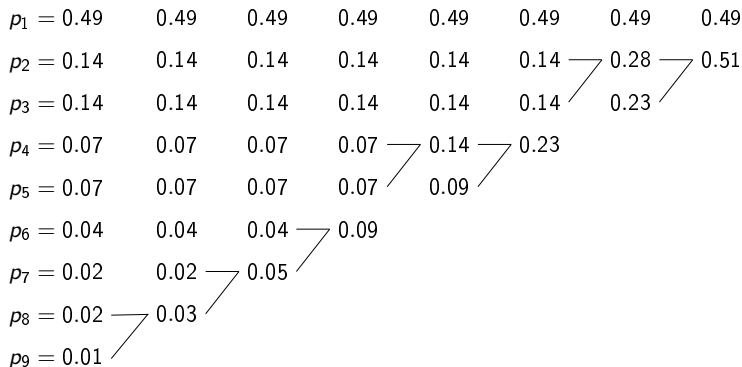
Megoldás. Először az állapotgráfot adjuk meg.



### 3. feladat

Kódoljuk a 2. feladat forrását Huffman kódolással.

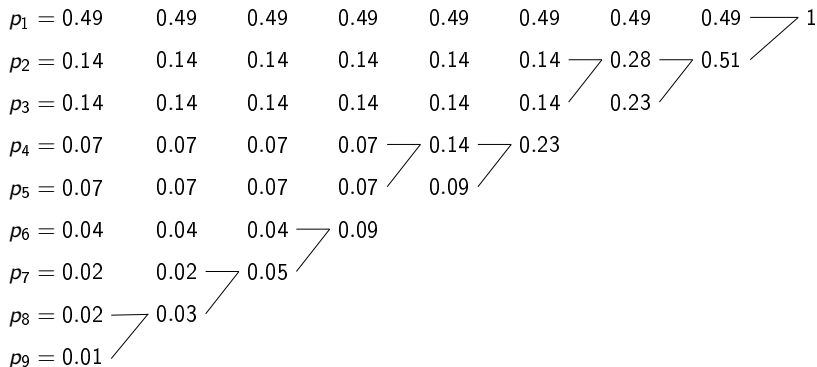
Megoldás. Először az állapotgráfot adjuk meg.



### 3. feladat

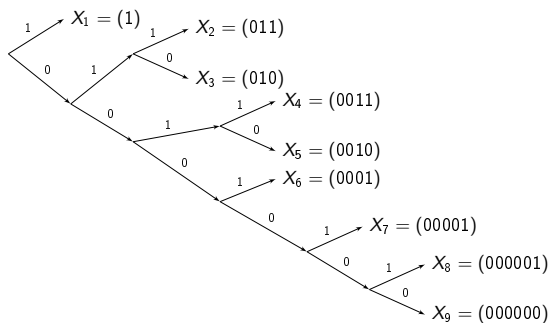
Kódoljuk a 2. feladat forrását Huffman kódolással.

Megoldás. Először az állapotgráfot adjuk meg.



## 4. feladat

Ezután a kódfa és a kódszavak táblázata (LUT) következik:



szimbólum	kódszó
$X_1$	1
$X_2$	011
$X_3$	010
$X_4$	0010
$X_5$	0011
$X_6$	0001
$X_7$	00001
$X_8$	000001
$X_9$	000000

## 4. feladat

Hasonlítsuk össze a Shannon-Fano kódolás és a Huffman kódolás teljesítményét az előző forrásra  $f_s = 160$  MHz mintavételi frekvencia esetén.

## 4. feladat

Hasonlítsuk össze a Shannon-Fano kódolás és a Huffman kódolás teljesítményét az előző forrásra  $f_s = 160$  MHz mintavételi frekvencia esetén.

Megoldás.

$$L^{HUFF} = 0.49 \cdot 1 + 0.14 \cdot 3 + 0.14 \cdot 3 + 0.07 \cdot 4 + 0.07 \cdot 4 + \\ + 0.04 \cdot 4 + 0.02 \cdot 5 + 0.02 \cdot 6 + 0.01 \cdot 6 = 2.33 \text{ (bit)}$$

$$L^{SF} = 0.49 \cdot 2 + 0.28 \cdot 3 + 0.14 \cdot 4 + 0.04 \cdot 5 + 0.04 \cdot 6 + \\ + 0.01 \cdot 7 = 2.89 \text{ (bit)}$$

Mivel  $f_s = 160$  MHz, a ráták

$$R_{HUFF} = 372.8 \text{ Mbps}, \quad R_{SF} = 462 \text{ Mbps}.$$

Megjegyzés: 9 forrás szimbólum  $\rightarrow$  adattömörítés nélkül 4 bitre van szükség, és a ráta  $R = 640 \text{ Mbps}$ .



## 5. feladat

Adott egy forrás a következő eloszlással és kódolással:

Forrás szimbólum	Valószínűség	Kódszó
$X_1$	0.4	0
$X_2$	0.2	10
$X_3$	0.2	110
$X_4$	0.2	1111

- (a) Mennyi az átlagos kódhossz?
- (b) Mennyi az elméleti alsó korlát a tömöríthetőségre?
- (c) Prefix-mentes-e ez a kód?
- (d) Optimális-e ez a kód?

## 5. feladat

Megoldás.

$$(a) L = \sum_{i=1}^4 p_i l_i = 0.4 \cdot 1 + 0.2 \cdot 2 + 0.2 \cdot 3 + 0.2 \cdot 4 = 2.2.$$

## 5. feladat

Megoldás.

(a)  $L = \sum_{i=1}^4 p_i l_i = 0.4 \cdot 1 + 0.2 \cdot 2 + 0.2 \cdot 3 + 0.2 \cdot 4 = 2.2.$

(b)

$$H(X) = \sum_{i=1}^4 p_i \log_2 \left( \frac{1}{p_i} \right) = 0.4 \cdot 1.31 + 3 \cdot 0.2 \cdot 2.322 = 1.922$$

$$L - H(X) = 0.278$$

## 5. feladat

Megoldás.

(a)  $L = \sum_{i=1}^4 p_i l_i = 0.4 \cdot 1 + 0.2 \cdot 2 + 0.2 \cdot 3 + 0.2 \cdot 4 = 2.2.$

(b)

$$H(X) = \sum_{i=1}^4 p_i \log_2 \left( \frac{1}{p_i} \right) = 0.4 \cdot 1.31 + 3 \cdot 0.2 \cdot 2.322 = 1.922$$

$$L - H(X) = 0.278$$

(c) Igen, a kód prefix-mentes.

## 5. feladat

Megoldás.

(a)  $L = \sum_{i=1}^4 p_i l_i = 0.4 \cdot 1 + 0.2 \cdot 2 + 0.2 \cdot 3 + 0.2 \cdot 4 = 2.2.$

(b)

$$H(X) = \sum_{i=1}^4 p_i \log_2 \left( \frac{1}{p_i} \right) = 0.4 \cdot 1.31 + 3 \cdot 0.2 \cdot 2.322 = 1.922$$

$$L - H(X) = 0.278$$

(c) Igen, a kód prefix-mentes.

(d) Nem, az  $X_4$  kódszavának elég 1111 helyett 111 is (és úgy egyébként pont a Huffman-kódolást kapjuk, ami már optimális).

## 6. feladat

Vegyük ismét az 1. feladat forrását:

$$p_1 = 0.49, \quad p_2 = 0.14, \quad p_3 = 0.14, \quad p_4 = 0.07, \quad p_5 = 0.07, \\ p_6 = 0.04, \quad p_7 = 0.02, \quad p_8 = 0.02, \quad p_9 = 0.01.$$

- (a) Tömörítsük a forrást Shannon-Fano-Elias kódolással.
- (b) Számítsuk ki az átlagos kódhosszt.
- (c) Hasonlítsuk össze a kód teljesítményét ugyanezen forrás Shannon-Fano kódolásával és Huffman kódolásával is  $f_s = 160$  MHz mintavételi frekvencia mellett.

## 6. feladat

Megoldás.

(a)

$i$	$p_i$	$F(i)$	$\bar{F}(i)$	bináris	$\ell_i$	kódszó
1	0.49	0	0.245	0.0011111010...	3	001
2	0.14	0.49	0.56	0.1000111101...	4	1000
3	0.14	0.63	0.7	0.1011001100...	4	1011
4	0.07	0.77	0.805	0.1100111000...	5	11001
5	0.07	0.84	0.875	0.1110000000...	5	11100
6	0.04	0.91	0.93	0.1110111000...	6	111011
7	0.02	0.95	0.96	0.1111010111...	7	1111010
8	0.02	0.97	0.98	0.1111101011...	7	1111101
9	0.01	0.99	0.995	0.1111111010...	8	11111110

$$F(i) = \sum_{j=0}^{i-1} p_j, \quad \bar{F}(i) = F(i) + p_i/2, \quad \ell_i = \lceil \log_2(1/p_i) \rceil + 1$$

## 6. feladat

(b) Az átlagos kódhossz

$$\begin{aligned}L^{SFE} &= 0.49 \cdot 3 + 0.14 \cdot 4 + 0.14 \cdot 4 + 0.07 \cdot 5 + 0.07 \cdot 5 + \\ &+ 0.04 \cdot 6 + 0.02 \cdot 7 + 0.02 \cdot 7 + 0.01 \cdot 8 = 3.89.\end{aligned}$$



## 6. feladat

(b) Az átlagos kódhossz

$$L^{SFE} = 0.49 \cdot 3 + 0.14 \cdot 4 + 0.14 \cdot 4 + 0.07 \cdot 5 + 0.07 \cdot 5 + \\ + 0.04 \cdot 6 + 0.02 \cdot 7 + 0.02 \cdot 7 + 0.01 \cdot 8 = 3.89.$$

(c)

$L^{HUFF} = 2.33$	$L^{SF} = 2.89$	$L^{SFE} = 3.89$
↓	↓	↓
$R_{HUFF} = 372.8Mbps$	$R_{SF} = 462Mbps$	$R_{SFE} = 622Mbps$

Emlékeztető: kódolás nélkül  $R = 640Mbps$ .

Konklúzió: az  $L$  átlagos kódhosszban elért kis javulás is sokat számít az adatsebességben!

# Összehasonlító elemzés

teljesítmény

$f_s = 160$  MHz

alg. egyszerűség

Kódolás	Teljesítmény	Átl. kódhossz	Adatseb.	Komplexitás
Huffman	$L$ optimális	2.33	372.4 Mbps	keresés + fa
SF	$H(X) < L < H(X) + 1$	2.89	462.4 Mbps	fa
SFE	$H(X) + 1 < L < H(X) + 2$	3.89	622.4 Mbps	bináris konverzió

## 7. feladat

Egy memóriamentes, időben homogén forrásnak a következő az eloszlása:

$$p_1 = 0.7, \quad p_2 = 0.2, \quad p_3 = 0.1.$$

- (a) Adjuk meg az elméleti alsó korlátot az átlagos kódhosszra.
- (b) Számítsuk ki az átlagos kódhosszt Huffman kódolás esetén.
- (c) Mennyi az átlagos kódhossz, ha a forrást Shannon-Fano kódolással tömörítjük?

## 7. feladat

Egy memóriamentes, időben homogén forrásnak a következő az eloszlása:

$$p_1 = 0.7, \quad p_2 = 0.2, \quad p_3 = 0.1.$$

- (a) Adjuk meg az elméleti alsó korlátot az átlagos kódhosszra.
- (b) Számítsuk ki az átlagos kódhosszt Huffman kódolás esetén.
- (c) Mennyi az átlagos kódhossz, ha a forrást Shannon-Fano kódolással tömörítjük?

## 7. feladat

Megoldás.

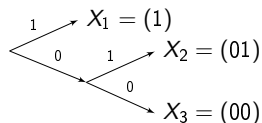
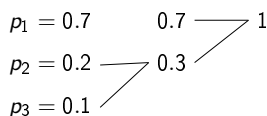
$$(a) H(X) = \sum_{i=1}^3 p_i \log_2(1/p_i) = 0.3598 + 0.464 + 0.3321 = 1.1559.$$

## 7. feladat

Megoldás.

(a)  $H(X) = \sum_{i=1}^3 p_i \log_2(1/p_i) = 0.3598 + 0.464 + 0.3321 =$   
**1.1559.**

(b)



$X_1$	1
$X_2$	01
$X_3$	00

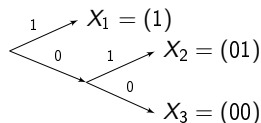
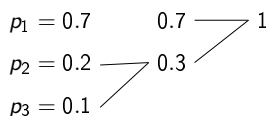
$$L_{HUFF} = 1 \cdot 0.7 + 2 \cdot 0.2 + 2 \cdot 0.1 = \mathbf{1.3}.$$

## 7. feladat

Megoldás.

(a)  $H(X) = \sum_{i=1}^3 p_i \log_2(1/p_i) = 0.3598 + 0.464 + 0.3321 =$   
**1.1559.**

(b)



$X_1$	1
$X_2$	01
$X_3$	00

$$L_{HUFF} = 1 \cdot 0.7 + 2 \cdot 0.2 + 2 \cdot 0.1 = \mathbf{1.3}.$$

(c)  $L_{SF} = \lceil \log_2(1/p_1) \rceil \cdot p_1 + \lceil \log_2(1/p_2) \rceil \cdot p_2 + \lceil \log_2(1/p_3) \rceil \cdot p_3 =$   
 $0.7 \cdot 1 + 0.2 \cdot 3 + 0.1 \cdot 4 = \mathbf{1.7}.$