

## II. Számpélda: bináris Huffman forráskódolás

Legyenek a diszkrét, hatelemű memóriamentes forrás (DMS) lehetséges szimbólumai (forrás ABC)  $X = \{A, B, C, D, E, F\}$  rendre a következő a-priori valószínűségekkel:

$$p(X) = \{p_A = 0.03125, p_B = 0.0625, p_C = 0.03125, p_D = 0.125, p_E = 0.25, p_F = 0.5\}.$$

- a.) Adja meg a forrás entrópiáját! (4 tizedesjegy pontossággal)
- b.) Adja meg az átlagos kódhosszt ( $L(X)$ )! (4 tizedesjegy pontossággal)
- c.) Adja meg a kód hatékonyságát!
- d.) Két szimbólumot összefogva, így a forráskiterjesztés módszerével növelve a forrás abc-t, hány %-kal tudjuk megnövelni a Huffman-kódoló hatékonyságát ( $h$ )?
- e.) Adja meg a legrövidebb bináris kód hosszát a szimbólumok számával!
- f.) Képezhető olyan kód az adott valószínűségek esetén, hogy a "01010" érvényes forráskód legyen?
- g.) Tegyük fel, hogy nincs pontos ismeretünk az alábbi  $p(X)$  forráseloszlásról, csak becsüljük azt  $q(X)$ -el a

$$p(X) = \{p_A = 0.03125, p_B = 0.03125, p_C = 0.0625, p_D = 0.125, p_E = 0.25, p_F = 0.5\}.$$

$q(X) = \{q_A = 0.03, q_B = 0.03, q_C = 0.06, q_D = 0.13, q_E = 0.25, q_F = 0.5\}$  valószínűségekkel. Szimbólumonként kódolva, hány %-kal változik a kódoló hatékonysága? (Ügyeljen arra, hogy a szimbólumok valószínűségi sorrendje nem változott  $q(X)$ -ben a  $p(X)$ -hez képest!)

a.) A forrás entrópiája [Shannon/szimbólum] (4 tizedesjegy pontossággal pl. 1,2345):

The correct answer is: 1.9375

b.) Az átlagos kódhossz [bit/szimbólum] (4 tizedesjegy pontossággal):

The correct answer is: 1.9375

c.) Kódhatékonyság [%]:

The correct answer is: 100

d.) A kódhatékonyság változása [%]:

The correct answer is: 0 %

e.) Adja meg a legrövidebb bináris kód hosszát a szimbólumok számával!

The correct answer is: 1

f.) Képezhető olyan kód az adott valószínűségek esetén, hogy a 01010 érvényes forráskód legyen?

Select one:

- a. igen ✓
- b. nem
- c. nem eldönthető

The correct answer is: igen

g.) A kódoló hatékonyságának változása  $q(X)$  feltételezésével [%]:

The correct answer is: 0 %

Adott egy diszkrét memóriamentes négy kimenetű forrás  $X$ , melynek ismert az a-priori első rendű eloszlása. Az aritmetikai kódolás alkalmazásával bevezetjük a 'STOP' szimbólumot  $p('STOP') = 1/16$  valószínűséggel. Az így adódó valószínűségi eloszlás az alábbi:

$$p(X) = [p(a)=1/4, p(b)=1/8, p(c)=1/2, p(d)=1/16, p('STOP')=1/16].$$

- a.) Adja meg az **első három** forrás szimbólumot, amelyet az 1101000001111..... kezdetű bináris kódsorozattal kódoltak el.
- b.) Adja meg a legrövidebb bináris kódsorozatot, amellyel a **DCB'STOP'** forrás sorozatot kódolja.

The correct answer is:

BAC

b.) Adja meg a legrövidebb bináris kódsorozatot, amellyel a forrás sorozatot kódolja! (pl. 11110001111)

The correct answer is:

111001101111

**Maximális távolságú (MDS), nemináris Hamming kód**

**A  $GF(q = 7)$  véges test felett,  $\alpha$  primitív elemmel számolva:**

határozza meg a maximális távolságú (MDS), perfekt Hamming kód  $(N, K, q)$  paramétereit, azaz adja meg  $N$  és  $K$  értékeit!

$N =$

The correct answer is: 8

$K =$

The correct answer is: 6

Határozza meg a lehetségesek közül azt a  $H$  mátrixot, melynél a nem egységmátrixot képző részben az egyik sorban az  $\alpha = 3$  primitív elem hatványai szerepelnek  $0$  kitevővel indulva, a kitevő szerint növekvő sorrendben  $(\alpha^0, \alpha^1, \alpha^2, \dots)$ , és olyan kódhoz tartozik, ahol a kódszó *eleje* azonos az üzenetszóval!

Adja meg a  $H$  mátrix *második* sorát, a szimbólumokat szóközők nélkül felsorolva!

Példa a megadási formátumra: 01230123

The correct answer is: 13264501

Határozza meg az ehhez a  $H = \begin{bmatrix} 11111110 \\ 13264501 \end{bmatrix}$  mátrixhoz tartozó  $G$  generátormátrix 4. sorát, a szimbólumokat szóközők nélkül felsorolva!

Példa a megadási formátumra: 01230123

The correct answer is: 00010061

Adja meg az ezzel a  $H = \begin{bmatrix} 11111110 \\ 15462301 \end{bmatrix}$  mátrixszal meghatározott  $s$  szindróma vektort (a szimbólumokat szóközők nélkül felsorolva) a  $v = [62142034]$  vett vektor esetére.

Példa a megadási formátumra: 01

The correct answer is: 43

Adja meg ezen  $H = \begin{bmatrix} 11111110 \\ 01326451 \end{bmatrix}$  mátrix és  $s = [42]$  szindróma vektor esetén az  $e$  hibavektort (a szimbólumokat szóközők nélkül felsorolva)!

Példa a megadási formátumra: 01230123

The correct answer is: 00000400

A  $v = [62142034]$  vett vektort véve a következő  $e = [00400000]$  hibavektort detektáltuk. Adja meg a legvalószínűbb  $c$  kódvektort (a szimbólumokat szóközők nélkül felsorolva)!

Példa a megadási formátumra: 01230123

The correct answer is: 62442034

Hibajavításkor ezzel a  $H = \begin{bmatrix} 11111110 \\ 01546231 \end{bmatrix}$  mátrixszal számolva az itt megadott  $c = [62442034]$  kódvektorra döntöttünk. Adja meg a legvalószínűbb  $u$  üzenetvektort (a szimbólumokat szóközők nélkül felsorolva)!

Példa a megadási formátumra: 012012

The correct answer is: 244203