

Vježbe 4

1. Data je poruka:

7 8 9 11 9 8 7 6 5 3 5 6 7 8 9 11 9 8 7 6 5 3 5 6 7

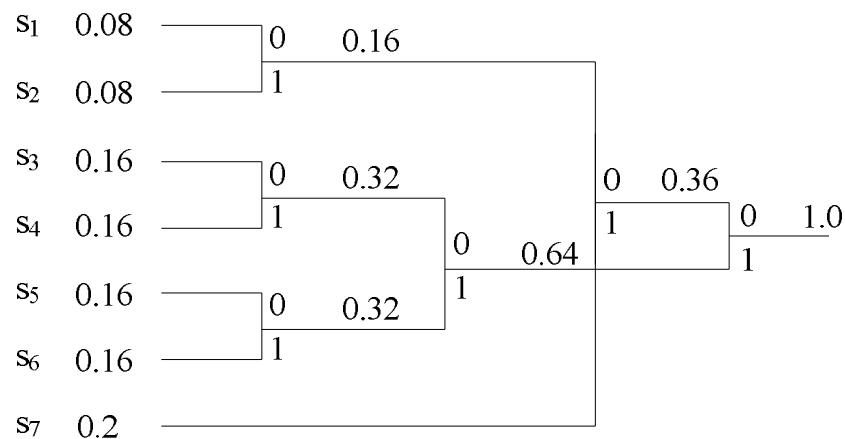
Kodirati ovakvu poruku Huffman-ovim kodom i odrediti prosječnu dužinu kodne riječi. Odrediti za svaki simbol odgovarajući kod. Odraditi diferencijalno kodiranje i primjeniti Huffman-ov kod na diferencijalnom kodu. Odrediti prosječnu dužinu kodne riječi i kodne simbole i u ovom slučaju.

Rješenje:

Ukupan broj simbola u datoj sekvenci je 25. Shodno tome, vjerovatnoće simbola koji se u njoj pojavljuju su:

$$p(3)=2/25=0.08; \quad p(5)=4/25=0.16; \quad p(6)=4/25=0.16; \quad p(7)=5/25=0.2;$$

$$p(8)=4/25=0.16; \quad p(9)=4/25=0.16; \quad p(11)=2/25=0.08;$$



s₁: 000 s₄: 101 s₇: 01
s₂: 001 s₅: 110
s₃: 100 s₆: 111

$$L = \sum_{i=1}^N l_i p_i = 3(0.08 + 0.08 + 0.16 + 0.16 + 0.16 + 0.16) + 2 * 0.2 = 1.2$$

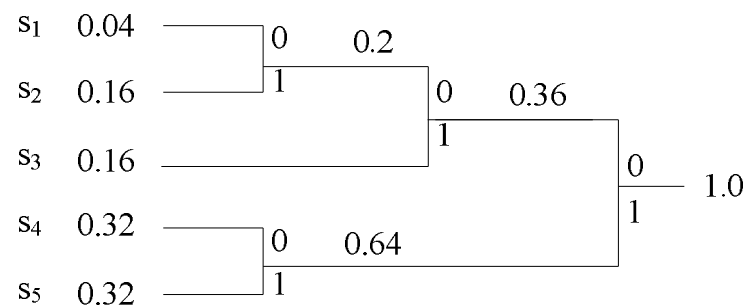
Diferencijalni kod je:

7 8 9 11 9 8 7 6 5 3 5 6 7 8 9 11 9 8 7 6 5 3 5 6 7
7 1 1 2 -2 -1 -1 -1 -1 -2 2 1 1 1 1 2 -2 -1 -1 -1 -1 -2 2 1 1

I diferencijalni kod ima ukupno 25 simbola, koji se pojavljuju sa sljedećim vjerovatnoćama:

$$p(-1)=8/25=0.32; \quad p(1)=8/25=0.32; \quad p(2)=4/25=0.16; \quad p(-2)=4/25=0.16;$$

$$p(7)=1/25=0.04;$$



s_1 : 000 s_4 : 10
 s_2 : 001 s_5 : 11
 s_3 : 01

$$L = \sum_{i=1}^N l_i p_i = 3(0.04 + 0.16) + 2(0.16 + 0.32 + 0.32) = 2.2$$

2. Sistem ima 9 kodnih simbola sa vjerovatnoćama: 0.23, 0.17, 0.16, 0.14, 0.12, 0.10, 0.03, 0.03, 0.02. Izvršiti kodiranje ternarnim (sa tri simbola) Huffman-ovim kodom. Izvršiti kodiranje Huffman-ovim kodom sa četiri simbola. Odrediti za svaki simbol odgovarajući kod, kao i prosječne dužine kodnih riječi u oba slučaja.

Rješenje:

a) Neka su $\{a, b, c\}$ simboli ternarnog Huffman-ovog koda. Očigledno, kod ovog Huffman-ovog koda se koristi nebinarni alfabet. U ovom slučaju je potrebno odrediti koliki broj kodnih simbola sa najmanjom vjerovatnoćom treba kodirati u prvom koraku. To se određuje na sljedeći način: od ukupnog broja kodnih poruka oduzima se po $D-1$ u svakom koraku (D je broj simbola u alfabetu Huffman-ovog koda). Broj kodnih poruka koji ostane u proceduri a manji je ili jednak D , a se kodira u prvom koraku:

$$D=3 \text{ i } Q=9$$

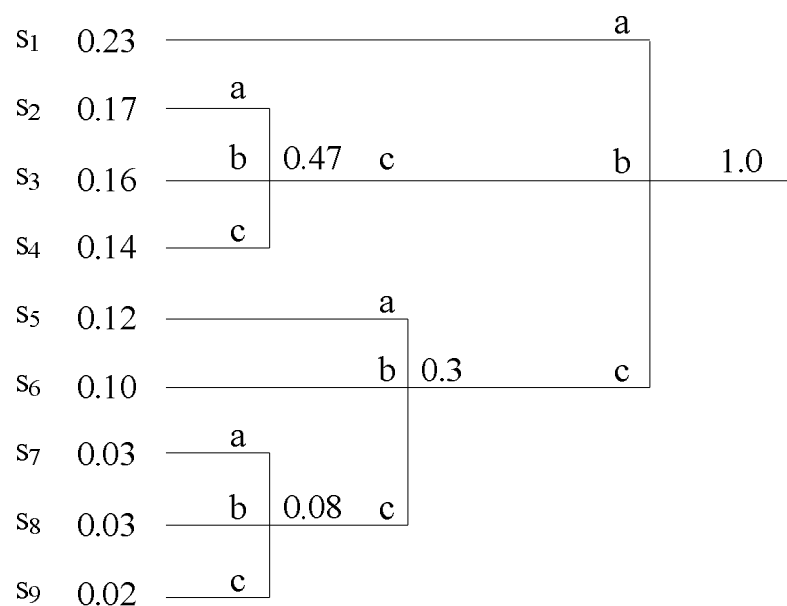
$$Q - (D-1) = Q$$

$$9 - 2 = 7$$

$$7 - 2 = 5$$

$$5 - 2 = 3$$

$3 \leq 3 \Rightarrow$ u prvom koraku uzimamo 3 kodne poruke



Kodne riječi su:

- s₁: a s₇: cca
- s₂: ba s₈: ccb
- s₃: bb s₉: ccc
- s₄: bc
- s₅: ca
- s₆: cb

$$L = \sum_{i=1}^N l_i p_i = 0.23 + 2(0.17 + 0.16 + 0.14 + 0.12 + 0.1) + 3(0.03 + 0.03 + 0.02) = 0.23 + 1.38 + 0.24 = 1.85$$

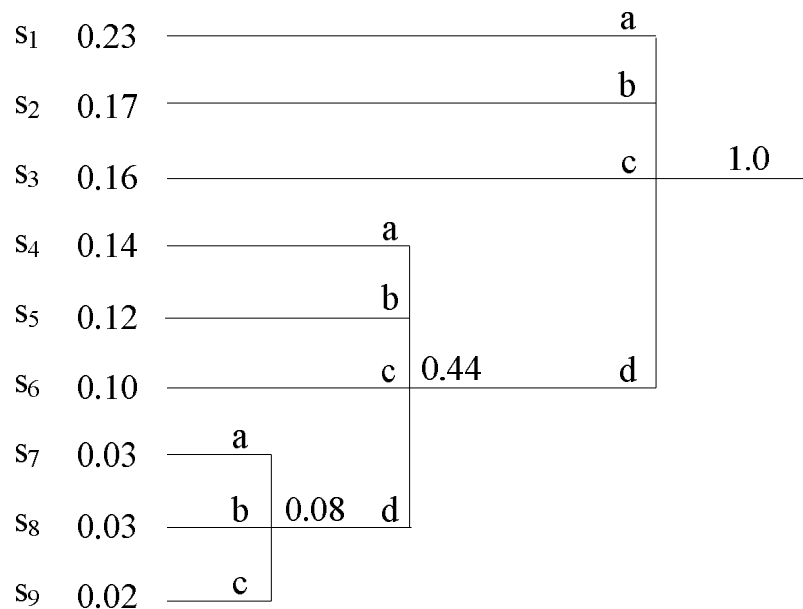
b) $D=4$ i $Q=9$

$$Q - (D - 1) = Q$$

$$9 - 3 = 6$$

$$6 - 3 = 3$$

$3 \leq 4 \Rightarrow$ u prvom koraku uzimamo 3 kodne poruke



Kodne riječi su:

- s₁: a s₇: dda
- s₂: b s₈: ddb
- s₃: c s₉: ddc
- s₄: da
- s₅: db
- s₆: dc

$$L = \sum_{i=1}^N l_i p_i = 0.23 + 0.17 + 0.16 + 2(0.14 + 0.12 + 0.1) + 3(0.03 + 0.03 + 0.02) = 0.56 + 0.72 + 0.24 = 1.52$$

3. Dat je string:

abacbabcabcbababcbacbbababaaabbcababa

Treba kodirati dati string LZ kodom. Rječnik ima 16 upisa i na početku sadrži kod izvora. Nakon popunjavanja, rječnik se resetuje.

Rješenje:

0	null
1	a
2	b
3	c
4	ab
5	ac
6	ba
7	bc
8	bab
9	cb
10	aba
11	bcb
12	acb
13	baba
14	aa
15	bb
0	null
1	a
2	b
3	c
4	cb
5	ab

(1,b) (1,c) (2,a) (2,c) (6,b) (3,b) (4,a) (7,b) (5,b) (8,a) (1,a) (2,b)

(3,b) (1,b) (1,null)