

ELEKTROTEHNIČKI FAKULTET UNIVERZITETA
SARAJEVO

Diskretna Matematika

Zadaća 3

Student

Vedad Fejzagić

Broj indeksa

17336

Demonstrator

Šeila Bećirović

Grupa

RI2-2

December 22, 2017

Zadatak 1

Postavka:

Neki eksperiment može dovesti do tri moguća događaja A_1 , A_2 ili A_3 iz skupa događaja X . Ova tri događaja imaju respektivno vjerovatnoće 0.25, 0.5 i 0.25. Rezultati tog eksperimenta nisu dostupni direktno, ali se može izvesti testni eksperiment koji daje događaje B_1 , B_2 , B_3 , B_4 ili B_5 iz skupa događaja Y , koji su u određenoj vezi sa događajima A_1 , A_2 i A_3 . Vjerovatnoće da testni eksperiment rezultira događajem B_j , $j = 1, 2, 3, 4, 5$ ukoliko je izvorni eksperiment rezultirao događajem A_i , $i = 1, 2, 3$ date su u sljedećoj tabeli:

$p(B_j / A_i)$	B_1	B_2	B_3	B_4	B_5
A_1	0.05	0.25	0.1	0.5	0.1
A_2	0.15	0.35	0.05	0.35	0.1
A_3	0.5	0.05	0.15	0.1	0.2

Odredite entropije skupa izvornih i testnih događaja $H(X)$ i $H(Y)$, uvjetne entropije $H(X/Y)$ i $H(Y/X)$, zajedničku entropiju $H(X,Y)$ te srednju količinu informacije $I(X,Y)$ koju testni događaji nose o izvornim događajima.

Rješenje:

$$H(X) = -\frac{1}{\ln 2} \cdot (0.25 \ln 0.25 + 0.5 \ln 0.5 + 0.25 \ln 0.25) = 1.5$$

Računamo vrijednosti $p(B_j)$ za $j = 1, 2, 3, 4, 5$:

$$p(B_j) = \sum_{i=1}^3 p(A_i) \cdot p(B_j/A_i)$$

Dobijamo:

$$p(B_1) = p(A_1)p(B_1/A_1) + p(A_2)p(B_1/A_2) + p(A_3)p(B_1/A_3) = 0.2125$$

$$p(B_2) = 0.25$$

$$p(B_3) = 0.0875$$

$$p(B_4) = 0.325$$

$$p(B_5) = 0.125$$

$$H(Y) = -\frac{1}{\ln 2} \cdot (0.2125 \ln 0.2125 + 0.25 \ln 0.25 + \\ + 0.0875 \ln 0.0875 + 0.325 \ln 0.325 + 0.125 \ln 0.125) = 2.178432$$

Sada računamo $p(A_i B_j)$, za $i = 1, 2, 3; j = 1, 2, 3, 4, 5$ Rezultati prikazani u vidu tabele:

$p(A_i B_j)$	B_1	B_2	B_3	B_4	B_5
A_1	0.0125	0.0625	0.025	0.125	0.025
A_2	0.075	0.175	0.025	0.175	0.05
A_3	0.125	0.0125	0.0375	0.025	0.05

$$H(X, Y) = - \sum_{i=1}^3 \sum_{j=1}^5 p(A_i B_j) \cdot \log_2 p(A_i B_j)$$

$$H(X, Y) = -\frac{1}{\ln 2} \sum_{i=1}^3 \sum_{j=1}^5 p(A_i B_j) \cdot \ln p(A_i B_j)$$

$$H(X, Y) = 3.4604$$

$$H(X/Y) = H(X, Y) - H(Y) = 1.27608$$

$$H(Y/X) = H(X, Y) - H(X) = 1.9604$$

$$I(X, Y) = H(X) + H(Y) - H(X, Y) = 0.218$$

Zadatak 2

Postavka:

Na nekom fakultetu, troškove studija za 22% studenata plaća država, dok su ostali studenti samofinansirajući. Među studentima koji se školuju o trošku države, 47% studenata stanuje u studentskom domu, dok među samofinansirajućim studentima 32% studenata stanuje u studentskom domu. Svi studenti koji stanuju u studentskom domu ujedno posjeduju i iskaznicu za subvencionirani javni prevoz, dok među studentima koji ne stanuju u studentskom domu istu iskaznicu posjeduju i 32% studenata čiji studij plaća država te 40% samofinansirajućih studenata.

Odredite koliku prosječnu količinu informacije saznanje o tome posjeduje li student iskaznicu za subvencionirani javni prenos ili ne nosi o načinu finansiranja njegovog studija (tj. da li ga finansira država ili troškove snosi sam).

Rješenje:

Zadatak 3

Postavka:

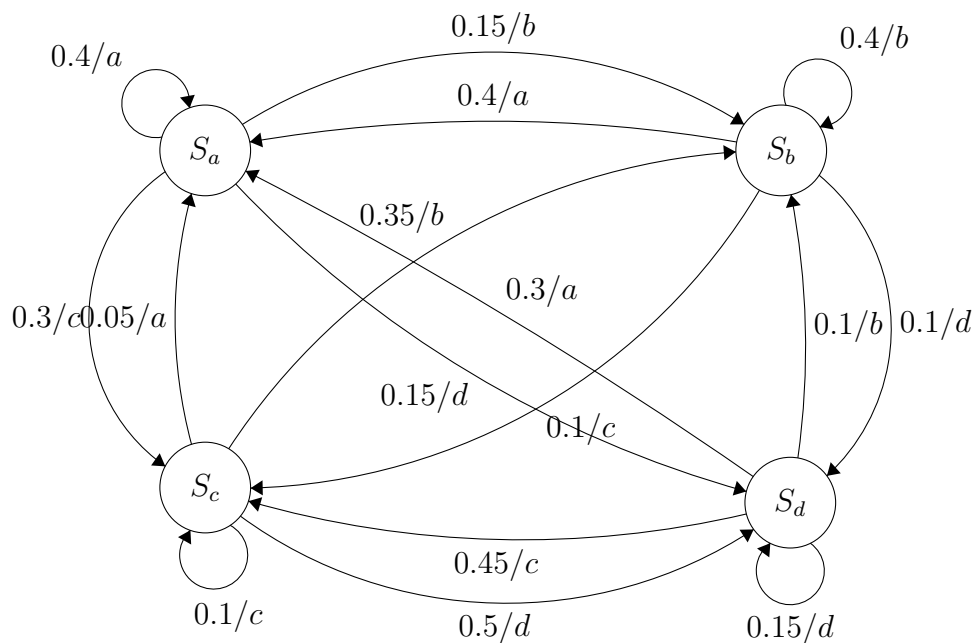
Markovljev izvor informacija prvog reda emitira četiri različite poruke a, b, c i d. Ovisno od toga koja je poruka posljednja emitirana, izvor se nalazi u jednom od 4 moguća stanja S_a , S_b , S_c i S_d koja redom odgovaraju emitiranim porukama a, b, c odnosno d. Vjerovatnoće da će izvor emitirati neku od ove 4 poruke ovisno od stanja u kojem se nalazi date su u sljedećoj tablici:

$p(x_j / S_i)$	a	b	c	d
S_a	0.4	0.15	0.3	0.15
S_b	0.4	0.4	0.1	0.1
S_c	0.05	0.35	0.1	0.5
S_d	0.3	0.1	0.45	0.15

Odredite entropiju i redudansu ovog izvora, zatim entropiju sekvenci dužine 6 te vjerovatnoću pojave sekvence aaabdb.

Rješenje:

Red izvora je $r = 1$, izvor modeliramo pomoću 4 stanja:



Računamo vjerovatnoće za svako od stanja rješavanjem sljedećeg sistema jednačina:

$$p(S_a) = p(S_a)p(a/S_a) + p(S_b)p(a/S_b) + p(S_c)p(a/S_c) + p(S_d)p(a/S_d)$$

$$p(S_b) = p(S_a)p(b/S_a) + p(S_b)p(b/S_b) + p(S_c)p(b/S_c) + p(S_d)p(b/S_d)$$

$$p(S_c) = p(S_a)p(c/S_a) + p(S_b)p(c/S_b) + p(S_c)p(c/S_c) + p(S_d)p(c/S_d)$$

$$p(S_a) + p(S_b) + p(S_c) + p(S_d) = 1$$

Nakon uvrštavanja vrijednosti i prebacivanja $p(S_a), p(S_b), p(S_c)$ na desnu stranu jednakosti, dobijamo sljedeću matricu:

$$M = \begin{bmatrix} -0.6 & 0.4 & 0.05 & 0.3 & |0 \\ 0.15 & -0.6 & 0.35 & 0.1 & |0 \\ 0.3 & 0.1 & -0.9 & 0.45 & |0 \\ 1 & 1 & 1 & 1 & |1 \end{bmatrix}$$

Koristimo Gausov metod eliminacije da riješimo zadani sistem, svođenjem matrice na desnu trougaonu matricu, dobijamo:

$$M = \begin{bmatrix} -0.6 & 0.4 & 0.05 & 0.3 & |0 \\ 0 & -0.5 & 0.36 & 0.17 & |0 \\ 0 & 0 & -0.65 & 0.70 & |0 \\ 0 & 0 & 0 & 4.54 & |1 \end{bmatrix}$$

$$-0.6p(S_a) + 0.4p(S_b) + 0.05p(S_c) + 0.3p(S_d) = 0$$

$$-0.5p(S_b) + 0.36p(S_c) + 0.17p(S_d) = 0$$

$$-0.65p(S_c) + 0.70p(S_d) = 0$$

$$4.54p(S_d) = 1$$

Dakle, rješenja sistema su:

$$p(S_a) = 0.29$$

$$p(S_b) = 0.25$$

$$p(S_c) = 0.23$$

$$p(S_d) = 0.22$$

$$H(S_a) = -\frac{1}{\ln 2}(p(a/S_a) \ln p(a/S_a) + p(b/S_a) \ln p(b/S_a) + p(c/S_a) \ln p(c/S_a) + p(d/S_a) \ln p(d/S_a))$$

Analogno za ostale, dobijamo:

$$H(S_a) \approx 1.871$$

$$H(S_b) \approx 1.722$$

$$H(S_c) \approx 1.578$$

$$H(S_d) \approx 1.782$$

$$H(X/X^\infty) = \sum_{i=1}^4 p(S_i) H(S_i) = 1.73$$

$$H_{max} = \log_2 4 = \frac{\ln 4}{\ln 2} = 2$$

$$R = \frac{H_{max} - H(X/X^\infty)}{H_{max}} \approx 0.135 = 13.5\%$$

$$H(X) = -\frac{1}{\ln 2}(p(a) \ln p(a) + p(b) \ln p(b) + p(c) \ln p(c) + p(d) \ln p(d)) \approx 1.98614$$

$$H(X^6) = H(X) + (6 - 1)H(X/X^\infty) \approx 10.636$$

$$p(aaabdb) = p(a)p(a/a)p(a/a)p(b/a)p(d/b)p(b/d) = 0.0000696 = 0.00696\%$$

Zadatak 4

Postavka:

Markovljev izvor informacija drugog reda emitira dvije različite poruke 0 i 1. Ovisno od toga koje su dvije poruke posljednje emitirane, izvor se može naći u jednom od 4 moguća stanja S_{00} , S_{01} , S_{10} odnosno S_{11} (recimo, ukoliko su posljednje dvije emitirane poruke 0 i 1 tim redom, izvor će se nalaziti u stanju S_{01}). Vjerovatnoće emitiranja poruke 0 u svakom od tih stanja iznose:

$$p(0/S_{00}) = 0.9$$

$$p(0/S_{01}) = 0.7$$

$$p(0/S_{10}) = 0.1$$

$$p(0/S_{11}) = 0.2$$

Odredite entropiju i redudansu ovog izvora, zatim entropiju sekvenci dužine 6 te vjerovatnoću pojave sekvence 00101100.

Rješenje:

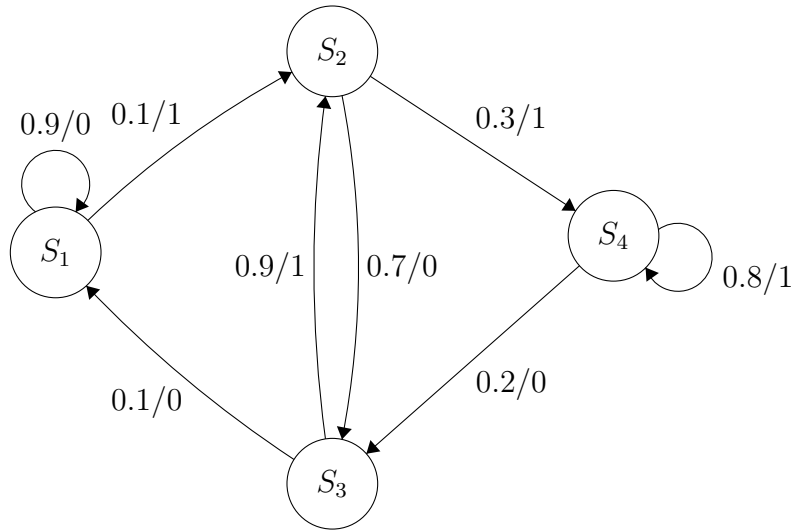
$$p(0/S_{00}) = 0.9, p(1/S_{00}) = 0.1$$

$$p(0/S_{01}) = 0.7, p(1/S_{01}) = 0.3$$

$$p(0/S_{10}) = 0.1, p(1/S_{10}) = 0.9$$

$$p(0/S_{11}) = 0.2, p(1/S_{11}) = 0.8$$

Neka su $S_{00} = S_1$, $S_{01} = S_2$, $S_{10} = S_3$, $S_{11} = S_4$.



Potrebne su nam vrijednosti $p(S_1), p(S_2), p(S_3), p(S_4)$

$$p(S_1) = p(S_1)p(0/S_1) + p(S_3)p(0/S_3)$$

$$0.1p(S_1) = 0.1p(S_3)$$

$$p(S_1) = p(S_3)$$

Analogno za ostale, dobije se:

$$p(S_2) = p(S_3)$$

$$0.3p(S_3) = 0.2p(S_4)$$

Potrebno je riješiti sljedeći sistem:

$$p(S_1) = p(S_3)$$

$$p(S_2) = p(S_3)$$

$$0.3p(S_3) = 0.2p(S_4)$$

$$p(S_1) + p(S_2) + p(S_3) + p(S_4) = 1$$

Rješavanjem sistema, dobiju se sljedeće vrijednosti:

$$p(S_1) = 0.2222$$

$$p(S_2) = 0.2222$$

$$p(S_3) = 0.2222$$

$$p(S_4) = 0.3333$$

$$H(S_1) = -\frac{1}{\ln 2}(p(0/S_1) \ln p(0/S_1) + p(1/S_1) \ln p(1/S_1)) \approx 0.47$$

$$H(S_2) \approx 0.881291$$

$$H(S_3) \approx 0.468996$$

$$H(S_4) \approx 0.721928$$

$$H(X/X^\infty) = \sum_{i=1}^4 p(S_i)H(S_i) = 0.638$$

Redudansa:

$$H_{max} = \log_2 4 = 2$$

$$R = \frac{H_{max} - H(X/X^\infty)}{H_{max}} \approx 0.681 \approx 68.1\%$$

Entropija sekvenci dužine 6:

$$H(X^2) = -\frac{1}{\ln 2}(p(00) \ln p(00) + p(01) \ln p(01) + p(10) \ln p(10) + p(11) \ln p(11)) = 1.96954$$

$$H(X^6) = H(X^2) + 4H(X/X^\infty) \approx 4.52154$$

Vjerovatnoća pojave sekvence 00101100:

$$p(00101100) = p(00)p(1/00)p(0/01)p(1/10)p(1/01)p(0/11)p(0/10) = 0.00008316 = 0.008316\%$$

Zadatak 5

Postavka:

Ergodični izvor informacija bez memorije emitira 10 poruka A, B, C, D, E, F, G, H, I i J. Proučavanjem sekvence dužine 645 koju je emitirao ovaj izvor, uočena je sljedeća učestalost pojavljivanja pojedinih poruka:

Poruka:	A	B	C	D	E	F	G	H	I	J
Učestalost:	18	73	87	49	73	99	98	44	80	24

Za ovaj izvor informacija formirajte:

- a) Binarni Shannon-Fano kod sa simbolima 0 i 1;
- b) Binarni Huffmanov kod sa simbolima 0 i 1;
- c) Ternarni Huffmanov kod sa simbolima 0, 1 i 2.

Za sva tri načina kodiranja, izračunajte protok informacija kroz komunikacioni kanal, procenat iskorištenja kanala veze, te kodirajte sekvencu poruka BAIBBBJIDJCE.

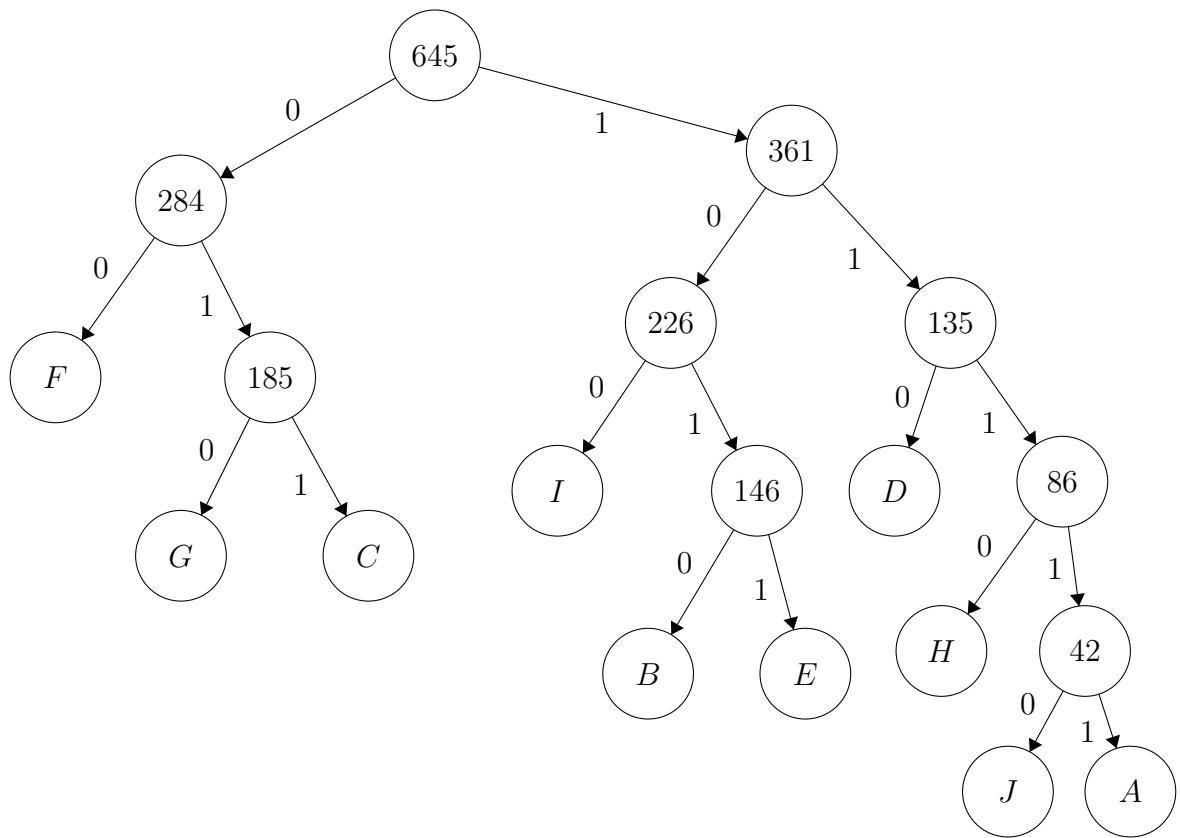
Rješenje:

a)

Sortiramo:

F	99
G	98
C	87
I	80
B	73
E	73
D	49
H	44
J	24
A	18

Kodiramo pomoću binarnog stabla:



Kodirane poruke su sljedeće:

F	00
G	010
C	011
I	100
B	1010
E	1011
D	110
H	1110
J	11110
A	11111

Kodirana sekvenca poruka: BAIBBBJIDJCE glasi:

1010111111001010101010101011110100110111100111011

Prosječna dužina kodne riječi:

$$n_{sr} = \sum_{i=1}^{10} p_i n_i \approx \sum_{i=1}^{10} \frac{N_i}{N} n_i = \frac{1}{N} \sum_{i=1}^{10} N_i n_i$$
$$n_{sr} = 3.271$$

Entropija izvora:

$$H(X/X^\infty) = H(X) = - \sum_{i=1}^{10} p_i \log_2 p_i = \dots = \frac{1}{\ln 2} (\ln N - \frac{1}{N} \sum_{i=1}^{10} N_i \ln N_i)$$
$$H(X/X^\infty) = 3.1671$$

Protok informacija:

$$\overline{I(H)} = \frac{H(X/X^\infty)}{n_{sr} \tau} \approx \frac{0.968}{\tau}$$

Iskorištenost kanala veze 96.8%

b)

Iz postavke: $n = 10, m = 2 \rightarrow m^* = 2 + \text{mod}(n - 4, m - 1) = 2$ Prelazimo na formiranje binarnog Huffmanovog koda:

Početak		Iteracija 1		Iteracija 2		Iteracija 3		Iteracija 4		Iteracija 5		Iteracija 6		Iteracija 7	
F	99	F	99	F	99	E/0	122	I/0	153	C/0	173	F/0	197	I/00	275
G	98	G	98	G	98	D/1		B/1		H/10		G/1		B/01	
C	87	C	87	C	87	F	99	E/0	122	J/110		C/0	173	E/10	
I	80	I	80	H/0	86	G	98	D/1		A/111		H/10		D/11	
B	73	B	73	J/10		C	87	F	99	I/0	153	J/110		F/0	197
E	73	E	73	A/11		H/0		G	98	B/1		A/111		G/1	
D	49	D	49	I	80	J/10	86	C	87	E/0	122	I/0	153	C/0	173
H	44	H	44	B	73	A/11		H/0	86	D/1		B/1		H/10	
J	24	J/0	42	E	73	I	80	J/10		F	99	E/0	122	J/110	
A	18	A/1		D	49	B	73	A/11		G	98	D/1		A/111	

Iteracija 8		Iteracija 9	
F/00	370	I/000	645
G/01		G/001	
C/10		C/010	
H/110		H/0110	
J/1110		J/01110	
A/1111		A/01111	
I/00	275	I/100	
B/01		B/101	
E/10		E/110	
D/11		D/111	

U sljedećim zadacima koji zahtijevaju veći broj iteracija, neće cijav postupak biti prikazan

$$n_{sr} = \frac{1}{645} \left(\sum_{i=1}^{10} N_i n_i \right) = 3.271$$

$$\overline{I(X)} = \frac{H(X/X^\infty)}{n_{sr}\tau} \approx \frac{0.968}{\tau}$$

Iskorištenost kanala veze 96.8%.

Kodirana sekvenca poruka: BAIBBBJIDJCE glasi:

101011111001011011010111010011101110010110

c)

Ternarni Huffmanov kod: $m = 0, m^* = 2 + \text{mod}(6, 2) = 2$

Početak		Iteracija 1		Iteracija 2		Iteracija 3		Iteracija 4		Iteracija 5	
F	99	F	99	D/0	135	I/0	226	F/0	284	F/00	645
G	98	G	98	H/1		B/1		G/1		G/01	
C	87	C	87	J/20		E/2		C/2		C/02	
I	80	I	80	A/21		D/0	135	I/0	226	I/10	
B	73	B	73	F	99	H/1		B/1		B/11	
E	73	E	73	G	98	J/20		E/2		E/12	
D	49	D	49	C	87	A/21		D/0	135	D/20	
H	44	H	44	I	80	F	99	H/1		H/21	
J	24	J/0	42	B	73	G	98	J/20		J/220	
A	18	A/1		E	73	C	87	A/21		A/221	

$$n_{sr} = \frac{1}{645} \left(\sum_{i=1}^{10} N_i n_i \right) = 2.061$$

$$\overline{I(X)} = \frac{H(X/X^\infty)}{n_{sr}\tau} \approx \frac{1.5367}{\tau}$$

Pošto je kapacitet ternarnog kanala veze $C_c = \frac{\log_2 3}{\tau} \approx \frac{1.5850}{\tau}$, iskorištenost je oko $\frac{1.5367}{1.5850} = 0.9695 = 96.95\%$

Kodirana sekvenca poruka: BAIBBBJIDJCE glasi:

112211011111122010202200212