

Zadaća 3

iz predmeta **Diskretna Matematika**

Prezime i ime: Hamzić Huso

Br. indexa: 18305

Asistent: Šeila Bećirović

Grupa: DM2 [Pon 15.00]

Zadatak	Bodovi
1	
2	
3	
4	
5	
6	
7	
8	
9	

Elektrotehnički fakultet Sarajevo

1. Rješenje zadatka

Zadatak 1 [0.25 poena]

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 0.2, 0.6 i 0.2. 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	0.15	0.2	0.2	0.4
A_2	0.05	0.15	0.05	0.35	0.4
A_3	0.35	0.2	0.2	0.15	0.1

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.

*Kako već imamo vjerovatnoće događaja X , entropiju $H(X)$ lahko računamo entropiju $H(X)$:

$$H(X) = - \sum_{i=1}^3 P(X_i) \cdot \log_2 (P(X_i))$$

odnosno:

$$H(X) = -(0.2 \cdot \log_2 0.2 + 0.6 \cdot \log_2 0.6 + 0.2 \cdot \log_2 0.2) = 1.37095$$

Kako bi odredili entropiju $H(Y)$ potrebne su nam sve vjerovatnoće $P(B_j)$ ($j = 1, 2, \dots, 5$) a njih lahko računamo :

$$P(B_1) = 0.2 \cdot 0.05 + 0.6 \cdot 0.05 + 0.2 \cdot 0.35 = 0.11 \quad P(B_2) = 0.2 \cdot 0.15 + 0.6 \cdot 0.15 + 0.2 \cdot 0.2 = 0.16$$

$$P(B_3) = 0.2 \cdot 0.2 + 0.6 \cdot 0.05 + 0.2 \cdot 0.2 = 0.11 \quad P(B_4) = 0.2 \cdot 0.2 + 0.6 \cdot 0.35 + 0.2 \cdot 0.15 = 0.28$$

$$P(B_5) = 0.2 \cdot 0.4 + 0.6 \cdot 0.4 + 0.2 \cdot 0.1 = 0.34$$

Sada možemo izračunati $H(Y)$:

$$H(Y) = - \sum_{j=1}^5 P(B_j) \cdot \log_2 (P(B_j))$$

odnosno:

$$H(Y) = -(0.11 \cdot \log_2 0.11 + 0.16 \cdot \log_2 0.16 + 0.11 \cdot \log_2 0.11 + 0.28 \cdot \log_2 0.28 + 0.34 \cdot \log_2 0.34)$$

odnosno:

$$\boxed{H(Y) = 2.166}$$

$H(X,Y)$ ćemo izračunati pomoću tabele, odnosno:

$$H(X,Y) = - \sum_{j=1}^5 \sum_{i=1}^3 P(B_j A_i) \cdot \log_2 (P(B_j A_i)) = 3.364$$

Sad ćemo izračunati $H(X/Y)$, $H(Y/X)$, $I(X,Y)$:

$$H(X/Y) = H(X,Y) - H(Y) = 3.364 - 2.166 = 1.198$$

$$H(Y/X) = H(X,Y) - H(X) = 3.364 - 1.37 = 1.994$$

$$I(X,Y) = H(X) - H(X/Y) = 1.37095 + 2.166 - 3.364 = 0.172$$

2. Rješenje zadatka

Zadatak 2 [0.25 poena]

Na nekom fakultetu, troškove studija za 33% 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 39% 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 posjeduje i 47% studenata čiji studij plaća država te 45% 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).

Skup sa događajima "troškove studija studenta plaća država" i "student je samofinansirajući" označimo sa X , pri čemu je prvi događaj X_1 a drugi događaj X_2 . Na osnovu teksta zadatka, poznato je da za ove događaje vrijedi da je

$$p(X_1) = 0.33 \quad p(X_2) = 0.67$$

Analogno neka je Y skup događaja gdje sada Y_1 predstavlja događaj "student stanuje u studentskom domu" a Y_2 "student ne stanuje u studentskom domu". Pa imamo uslovne vjerovatnoće:

$$p(Y_1/X_1) = 0.47 \quad p(Y_1/X_2) = 0.39$$

Moraju biti zadovoljene sljedeće jednakosti jer student kojeg država financira može ili da živi ili da ne živi u studentskom domu(isto ovo vrijedi i za samofinansirajuće studente). Pa imamo jednakosti:

$$p(Y_1/X_1) + p(Y_2/X_1) = 1 \quad p(Y_1/X_2) + p(Y_2/X_2) = 1$$

na osnovu čega se mogu izračunati sljedeće uslovne vjerovatnoće

$$p(Y_2/X_1) = 1 - p(Y_1/X_1) = 0.53 \quad p(Y_2/X_2) = 1 - p(Y_1/X_2) = 0.61$$

Neka je sad Z skup događaja koji sadrži događaje Z_1 - "student posjeduje iskaznicu za javni prevoz" i Z_2 - "student ne posjeduje iskaznicu za javni prevoz". Pa možemo odrediti vjerovatnoće:

$p(Z_1/X_1) = 0.47 + 0.47 \cdot (1 - 0.47) = 0.7191$. Jasno je da svaki student čije troškove studija plaća država ili ima ili nema iskaznicu, pa je vjerovatnoća $p(Z_2/X_1) = 1 - 0.7191 = 0.2809$

Na sličan način računa se vjerovatnoća da student posjeduje iskaznicu uz uslov da je on samofinansirajući, te ona iznosi $0.39 + 0.45 \cdot (1 - 0.39) = 0.6645$. Ovo je naša uvjetna vjerovatnoća $p(Z_1/X_2)$, odnosno njena suprotna vjerovatnoća je $p(Z_2/X_2) = 1 - 0.6645 = 0.3355$. Formirajmo tabelu za dobijene vrijednosti

$p(Z_j/X_i)$	Z_1	Z_2
X_1	0.7191	0.2809
X_2	0.6645	0.3355

Prosječna količina informacija koju informacija da li student posjeduje iskaznicu za subvencionirani javni prevoz ili ne nosi o načinu finansiranja njegovog studija označava se sa $I(X, Z)$ i računa se kao:

$$I(X, Z) = H(X) + H(Z) - H(X, Z)$$

Nađimo $H(X)$ kao:

$$H(X) = -(p(X_1) \cdot \log_2 p(X_1) + p(X_2) \cdot \log_2 p(X_2)) = -(0.33 \cdot \log_2 0.33 + 0.67 \cdot \log_2 0.67) = 0.91492$$

Analogno $H(Z)$:

$$H(Z) = -(p(Z_1) \cdot \log_2 p(Z_1) + p(Z_2) \cdot \log_2 p(Z_2))$$

Ali vjerovatnoće $p(Z_1)$ i $p(Z_2)$ nisu poznate, međutim lahko ih računamo:

$$p(Z_1) = p(X_1) \cdot p(Z_1/X_1) + p(X_2) \cdot p(Z_1/X_2) = 0.33 \cdot 0.7191 + 0.67 \cdot 0.6645 = 0.6825$$

$$p(Z_2) = 1 - p(Z_1) = 0.3175$$

Entropiju $H(Z)$ sada možemo izračunati:

$$H(Z) = -(0.6825 \cdot \log_2 0.6825 + 0.3175 \cdot \log_2 0.3175) = 0.901$$

Izračunajmo $H(X, Z)$:

$$H(X, Z) = -(p(X_1 Z_1) \cdot \log_2 p(X_1 Z_1) + p(X_2 Z_1) \cdot \log_2 p(X_2 Z_1) +$$

$$p(X_1 Z_2) \cdot \log_2 p(X_1 Z_2) + p(X_2 Z_2) \cdot \log_2 p(X_2 Z_2))$$

$$H(X, Z) = -(0.33 \cdot 0.7191 \cdot \log_2 (0.33 \cdot 0.7191) + 0.33 \cdot 0.2809 \cdot \log_2 (0.33 \cdot 0.2809) + \\ + 0.67 \cdot 0.6645 \cdot \log_2 (0.67 \cdot 0.6645) + 0.67 \cdot 0.3355 \cdot \log_2 (0.67 \cdot 0.3355)) = 1.8137$$

I na kraju imamo:

$I(X, Z) = H(X) + H(Z) - H(X, Z) = 0.91492 + 0.901 - 1.8137 = 0.00222$
--

3. Rješenje zadatka

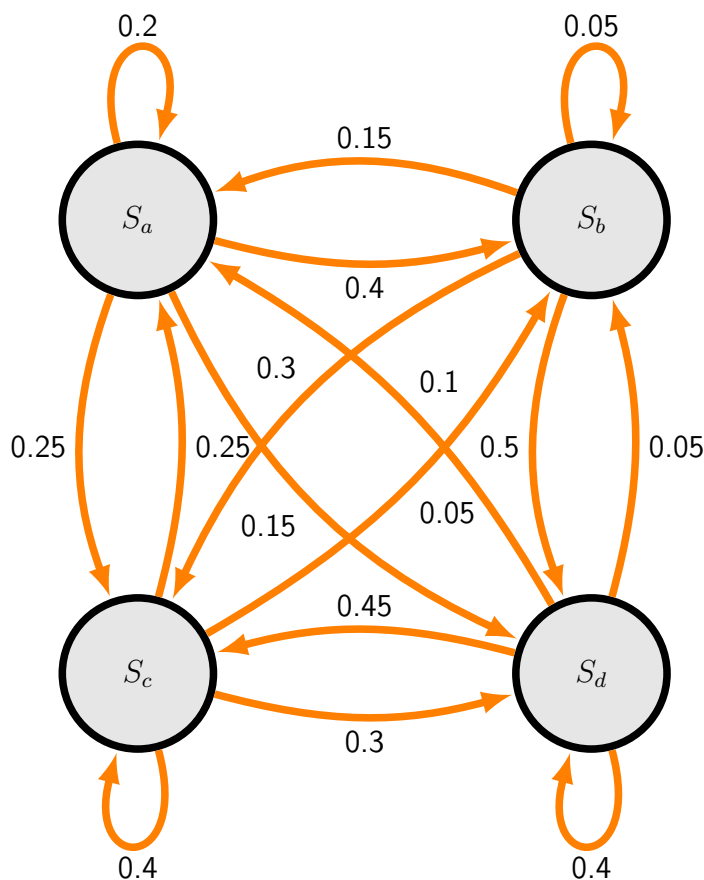
Zadatak 3 [0.35 poena]

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	0.2	0.4	0.25	0.15
S_b	0.15	0.05	0.3	0.5
S_c	0.25	0.05	0.4	0.3
S_d	0.1	0.05	0.45	0.4

Odredite entropiju i redudansu ovog izvora, zatim entropiju sekvenci dužine 7 te vjerovatnoću pojave sekvence bcbdbbc.

Izvor možemo modelirati pomoću 4 stanja, S_a , S_b , S_c , S_d koja redom odgovaraju prethodno emitiranim porukama a, b, c, d. Grafički prikaz konačnog automata koji modelira ovaj izvor dat je na slici ispod:



Stanje S_a može nastati prelazom iz stanja S_a, S_b, S_c ili S_d svaki put uz emitiranje poruke a. Na osnovu teoreme o totalnoj vjerovatnoći za svako stanje, imamo sisteme:

$$\begin{aligned}P(S_a) &= 0.2 \cdot P(S_a) + 0.15 \cdot P(S_b) + 0.25 \cdot P(S_c) + 0.1 \cdot P(S_d) \\P(S_b) &= 0.4 \cdot P(S_a) + 0.05 \cdot P(S_b) + 0.05 \cdot P(S_c) + 0.05 \cdot P(S_d) \\P(S_c) &= 0.25 \cdot P(S_a) + 0.3 \cdot P(S_b) + 0.4 \cdot P(S_c) + 0.45 \cdot P(S_d) \\P(S_a) + P(S_b) + P(S_c) + P(S_d) &= 1\end{aligned}$$

Rješenja našeg sistema su respektivno:

$$\boxed{P(S_a) = 0.18} \quad \boxed{P(S_b) = 0.113} \quad \boxed{P(S_c) = 0.378} \quad \boxed{P(S_d) = 0.3284}$$

Da bi odredili entropiju izvora $H(X/X^\infty)$ moramo odrediti entropije pojedinih stanja, odnosno $H(S_a), H(S_b), H(S_c), H(S_d)$:

$$\begin{aligned}H(S_a) &= -(P(a/S_a) \cdot \log_2 P(a/S_a) + P(b/S_a) \cdot \log_2 P(b/S_a) + \\&\quad P(c/S_a) \cdot \log_2 P(c/S_a) + P(d/S_a) \cdot \log_2 P(d/S_a))\end{aligned}$$

analogno je i za ostala stanja pa ćemo napisati samo rješenja jer se ovo uz pomoć grafika lahko računa

$$\boxed{H(S_a) = 1.9037} \quad \boxed{H(S_b) = 1.6477} \quad \boxed{H(S_c) = 1.7659} \quad \boxed{H(S_d) = 1.5954}$$

Za entropiju izraza imamo:

$$H(X/X^\infty) = P(S_a)H(S_a) + P(S_b)H(S_b) + P(S_c)H(S_c) + P(S_d)H(S_d) = 1.7202$$

Redudansu izvora računamo kao:

$$R = \frac{\log_2 4 - H(X/X^\infty)}{\log_2 4} = \frac{2 - 1.7202}{2} = 0.1399$$

Vjerovatnoću sekvence **bcbcdbbc** možemo izračunati kao

$$P(bcbcdbbc) = P(b) \cdot P(c/b) \cdot P(b/c) \cdot P(c/b) \cdot P(d/c) \cdot P(b/d) \cdot P(b/b) \cdot P(c/b)$$

odnosno

$$0.113 \cdot 0.3 \cdot 0.05 \cdot 0.3 \cdot 0.3 \cdot 0.05 \cdot 0.05 \cdot 0.3 = 1.144 \cdot 10^{-7}$$

Za računanje entropije sekvenci dužine 7 imamo relaciju:

$$H(X^7) = H(X^1) + 6 \cdot H(X/X^\infty)$$

Treba nam još entropija $H(X^1)$:

$$H(X^1) = -(P(a) \cdot \log_2 P(a) + P(b) \cdot \log_2 P(b) + P(c) \cdot \log_2 P(c) + P(d) \cdot \log_2 P(d)) = 1.8588$$

Izračunajmo sada entropiju sekvenci 7:

$$H(X^7) = H(X^1) + 6 \cdot H(X/X^\infty) = 1.8588 + 6 \cdot 1.7202 = 12.18$$

4. Rješenje zadatka

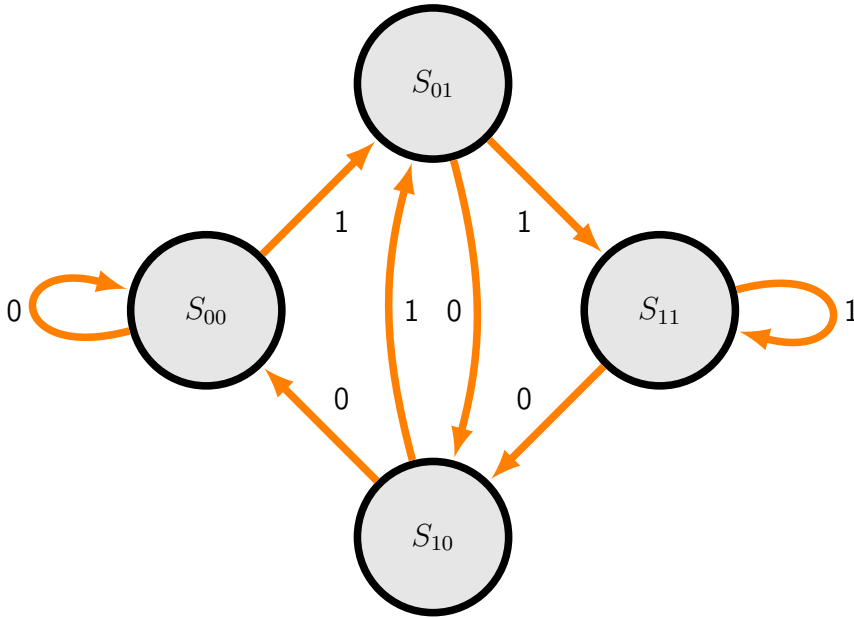
Zadatak 4 [0.4 poena]

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:

$$p(0/S_{00}) = 0.4 \quad p(0/S_{01}) = 0.3 \quad p(0/S_{10}) = 0.7 \quad p(0/S_{11}) = 0.1$$

Odredite entropiju i redudansu ovog izvora, zatim entropiju sekvenci dužine 5 te vjerovatnoću pojave sekvence 111001.

* Kako je red izvora $r = 2$, izvor možemo modelirati pomoću 4 stanja S_{00} , S_{01} , S_{10} odnosno S_{11} . Grafički prikaz konačnog automata koji modelira ovaj izvor prikazan je na slici ispod:



Kako imamo vjerovatnoće emitiranja poruke 0 u svim stanjima, na osnovu njih možemo dobiti vjerovatnoće emitiranja poruke 1 u tim istim stanjima:

$$p(1/S_{00}) = 0.6 \quad p(1/S_{01}) = 0.7 \quad p(1/S_{10}) = 0.3 \quad p(1/S_{11}) = 0.9$$

Izračunajmo entropiju stanja:

$$H(S_{00}) = -(0.4 \cdot \log_2 0.4 + 0.6 \cdot \log_2 0.6) = 0.9709$$

$$H(S_{01}) = -(0.3 \cdot \log_2 0.3 + 0.7 \cdot \log_2 0.7) = 0.881$$

$$H(S_{10}) = -(0.7 \cdot \log_2 0.7 + 0.3 \cdot \log_2 0.3) = 0.881$$

$$H(S_{11}) = -(0.1 \cdot \log_2 0.1 + 0.9 \cdot \log_2 0.9) = 0.4689$$

Sada možemo postaviti sistem jednačina pomoću kojeg dobijamo vjerovatnoće svakog od stanja:

$$\begin{aligned}P(S_{00}) &= 0.4 \cdot P(S_{00}) + 0.7 \cdot P(S_{10}) \\P(S_{01}) &= 0.6 \cdot P(S_{00}) + 0.3 \cdot P(S_{10}) \\P(S_{11}) &= 0.7 \cdot P(S_{01}) + 0.9 \cdot P(S_{11}) \\P(S_{00}) + P(S_{01}) + P(S_{10}) + P(S_{11}) &= 1\end{aligned}$$

Respektivno, rješenja su:

$$P(S_{00}) = 0.1147 \quad P(S_{01}) = 0.0983 \quad P(S_{10}) = 0.0983 \quad P(S_{11}) = 0.6885$$

Izračunajmo sada entropiju izvora

$$H(X/X^\infty) = P(S_{00})H(S_{00}) + P(S_{01})H(S_{01}) + P(S_{10})H(S_{10}) + P(S_{11})H(S_{11}) = 0.6074$$

Redudansa izvora je:

$$R = \frac{\log_2 2 - 0.6074}{1} = 0.39259$$

Odredimo entropiju sekvenci dužine 5:

$$H(X^5) = H(X^2) + 3 \cdot H(X/X^\infty)$$

Odredimo $H(X^2)$:

$$H(X^2) = -(P(S_{00}) \cdot \log_2 P(S_{00}) + P(S_{01}) \cdot \log_2 P(S_{01}) + P(S_{10}) \cdot \log_2 P(S_{10}) + P(S_{11}) \cdot \log_2 P(S_{11}))$$

odnosno

$$H(X^2) = 1.38702$$

odnosno entropija sekvenci 5 je:

$$H(X^5) = H(X^2) + 3 \cdot H(X/X^\infty) = 1.38702 + 3 \cdot 0.6074 = 3.20922$$

Izračunajmo još i vjerovatnoću sekvence 111001:

$$p(111001) = p(11) \cdot p(1/11) \cdot p(0/11) \cdot p(0/10) \cdot p(1/00) = 0.6885 \cdot 0.9 \cdot 0.1 \cdot 0.7 \cdot 0.6 = 0.0260253$$

odnosno

$$\boxed{p(111001) = 2.60253\%}$$

5. Rješenje zadatka

Zadatak 5 [0.6 poena]

Ergodični izvor informacija bez memorije emitira 10 poruka A, B, C, D, E, F, G, H, I i J. Proučavanjem sekvence dužine 449 koju je emitirao ovaj izvor, uočena je sljedeća učestalost pojavljivanja pojedinih poruka:

Poruka:	A	B	C	D	E	F	G	H	I	J
Učestalost:	39	77	30	69	49	25	82	28	10	40

Za ovaj izvor informacija formirajte:

- Binarni Shannon-Fano kod sa simbolima 0 i 1;
- Binarni Huffmanov kod sa simbolima 0 i 1;
- 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 BFHIAEDIB.

* a) Konstrukcija Shannon-Fano koda prikazana je u sljedećoj tabeli, pri čemu su u tabeli umjesto vjerovatnoća prikazane učestalosti, što se zapravo svodi na isto, jer su učestalosti proporcionalne vjerovatnoćama.

82 G	228/0	82/00		
77 B		146/01	77/010	
69 D			69/011	
49 E	221/1	128/10	49/100	
40 J			79/101	40/1010
39 A				39/1011
30 C		93/11	58/110	30/1100
28 H				28/1101
25 F			35/111	25/1110
10 I				10/1111

Imamo kodirane sekvence:

A-1011 B-010 C-1100 D-011 E-100 F-1110 G-00 H-1101 I-1111 J-1010

Izračunajmo $H(X/X^\infty)$ pri čemu sad moramo uzeti vjerovatnoće u obzir, odnosno učestalost / 449:

$$H(X/X^\infty) = -\frac{1}{449} \left(82 \cdot \log_2 \frac{82}{449} + 77 \cdot \log_2 \frac{77}{449} + 69 \cdot \log_2 \frac{69}{449} + 49 \cdot \log_2 \frac{49}{449} + 40 \cdot \log_2 \frac{40}{449} + \right. \\ \left. 39 \cdot \log_2 \frac{39}{449} + 30 \cdot \log_2 \frac{30}{449} + 28 \cdot \log_2 \frac{28}{449} + 25 \cdot \log_2 \frac{25}{449} + 10 \cdot \log_2 \frac{10}{449} \right)$$

odnosno:

$$H(X/X^\infty) = 3.1281$$

Da bi izračunali prosječnu dužinu kodne riječi, ukoliko je n_i dužina kodne riječi pridružene i-toj poruci koristimo formulu:

$$n_{sr} = \frac{1}{N} \cdot \sum_{i=1}^m N_i \cdot n_i$$

odnosno:

$$n_{sr} = \frac{1}{449} (82 \cdot 2 + 77 \cdot 3 + 69 \cdot 3 + 49 \cdot 3 + 40 \cdot 4 + 39 \cdot 4 + 30 \cdot 4 + 28 \cdot 4 + 25 \cdot 4 + 10 \cdot 4) = 3.200$$

pa je protok informacija:

$$\overline{I(X)} = \frac{H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{0.97753125}{\tau}$$

odakle slijedi da je iskorištenost kanala veze približno 97.753%

Kodirana poruka BFHIAEDIB glasi (razmak između svakog slova):
010 1110 1101 1111 1011 100 011 1111 010

b) Binarni Huffmanov kod sa 0 i 1:

82 G	82 G	82 G	82 G	E/0	D/0	B/0	E/00	B/00	B/000
77 B	77 B	77 B	77 B	89 J/1	C/10 127	A/10 151	J/01 171	A/010	A/0010
69 D	69 D	69 D	A/0	82 G	H/11	F/110	G/1	F/0110	F/00110
49 E	49 E	C/0	F/10 74	77 B	E/0	I/111	B/0	I/0111	I/00111
40 J	40 J	58 H/1	I/11	A/0	89 J/1	D/0	A/10	D/10	D/010
39 A	39 A	49 E	69 D	F/10 74	82 G	C/10 127	F/110	C/110	C/0110
30 C	F/0	40 J	C/0	I/11	77 B	H/11	I/111	H/111	H/0111
28 H	35 I/1	39 A	58 H/1	69 D	A/0	E/0	D/0	E/00	E/100
25 F	30 C	F/0	49 E	C/0	F/10 74	89 J/1	C/10 127	J/01	J/101
10 I	28 H	I/1	40 J	H/1	I/11	82 G	H/11	G/1	G/11

Po formuli koju smo koristili u dijelu zadatka pod a, dobijamo da je:

$$n_{sr} = \frac{1}{449}(3 \cdot 77 + 4 \cdot 39 + 5 \cdot 25 + 5 \cdot 10 + 3 \cdot 69 + 4 \cdot 30 + 4 \cdot 28 + 3 \cdot 49 + 3 \cdot 40 + 2 \cdot 82) = 3.189$$

pa je na osnovu toga protok informacija (pošto je entropija izvora ista kako pod a)

$$\overline{I(X)} = \frac{H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{0.9809}{\tau}$$

odakle slijedi da je iskorištenost kanala veze približno 98.09%

Kodirana poruka BFHIAEDIB glasi (razmak između svakog slova):
000 00110 0111 00111 0010 100 010 00111 000

c) Ternarni Huffmanov kod sa 0, 1 i 2 :

82 G	82 G	F/00	E/0	G/0	G/00
77 B	77 B	I/01	J/1 128	B/1 228	B/01
69 D	69 D	93 C/1	A/2	D/2	D/02
49 E	49 E	H/2	F/00	E/0	E/10
40 J	40 J	82 G	I/01	J/1 128	J/11
39 A	39 A	77 B	C/1	A/2	A/12
30 C	F/0 35	69 D	H/2	F/00	F/200
28 H	I/1	49 E	82 G	I/01 93	I/201
25 F	30 C	40 J	77 B	C/1	C/21
10 I	28 H	39 A	69 D	H/2	H/22

Na isti način kao u dijelovima zadatka pod a i b dobijamo da je:

$$n_{sr} = \frac{1}{449}(2 \cdot 82 + 2 \cdot 77 + 2 \cdot 69 + 2 \cdot 49 + 2 \cdot 40 + 2 \cdot 39 + 3 \cdot 25 + 3 \cdot 10 + 2 \cdot 30 + 2 \cdot 28) = 2.0779$$

Entropija nam je ista kao i pod a odnosno b, jer je isti skup podataka, pa na osnovu toga imamo protok informacija:

$$\overline{I(X)} = \frac{H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{1.5054}{\tau}$$

Kako je kapacitet kanala veze $C_c = \frac{\log_2 3}{\tau} = \frac{1.5850}{\tau}$ iskorištenost kanala veze je $\frac{1.5054}{1.5850} = 0.9498 = 94.98\%$

Kodirana poruka BFHIAEDIB glasi (razmak između svakog slova):
01 200 22 201 12 10 02 201 01

6. Rješenje zadatka

Zadatak 6 [0.7 poena]

Izvor informacija bez memorije emitira 4 poruke A, B, C i D. Vjerovatnoće pojavljivanja ovih poruka iznose:

$$\begin{aligned} p(A) &= 0.25 \\ p(B) &= 0.05 \\ p(C) &= 0.45 \\ p(D) &= 0.25 \end{aligned}$$

Za ovaj izvor informacija formirajte

- Binarni Shannon-Fano kod sa simbolima 0 i 1,
- Binarni Huffmanov kod sa simbolima 0 i 1,
- Binarni Shannon-Fano kod sa simbolima 0 i 1, ali kodirajući parove poruka umjesto individualnih poruka,
- Binarni Huffmanov kod sa simbolima 0 i 1, ali kodirajući parove poruka umjesto individualnih poruka.

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

* a) Binarni Shannon-Fano kod sa simbolima 0 i 1:

C 0.45	0.45/0		
A 0.25	0.55/1	0.25/10	
D 0.25		0.3/11	0.25/110
B 0.05			0.05/111

Iz tabele se vidi koja je poruka kodirana kojim kodom, izračunajmo entropiju izvora i prosječnu dužinu kodne riječi:

$$H(X/X^\infty) = -(0.45 \cdot \log_2 0.45 + 0.25 \cdot \log_2 0.25 + 0.25 \cdot \log_2 0.25 + 0.05 \cdot \log_2 0.05) = 1.73449$$

$$n_{sr} = 0.45 \cdot 1 + 0.25 \cdot 2 + 0.25 \cdot 3 + 0.05 \cdot 3 = 1.85$$

pa je protok kanala veze:

$$\overline{I(X)} = \frac{H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{0.93756}{\tau}$$

odnosno iskorištenost kanala veze je 93.756%.

Kodirana poruka BADCDBCBCB glasi (razmak između svakog slova):

111 10 110 0 110 111 0 111 0 111

b) Binarni Huffmanov kod sa simbolima 0 i 1:

C 0.45	C 0.45	D/00	D/000
A 0.25	D/0 0.3	B/01 0.55	B/001
D 0.25	B/1	A/1	A/01
B 0.05	A 0.25	C 0.45	C/1

Iz tabele se može vidjeti kojim su sekvencama slova kodirana. Entropija je ista kao u dijelu zadatka pod a, a isti je slučaj i sa prosječnom dužinom kodne riječi $n_{sr} = 1.85$, pa vrijede isti zaključci osim kodirane kodirane sekvence. Odnosno protok kanala veze je:

$$\overline{I(X)} = \frac{H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{0.93756}{\tau}$$

odnosno iskorištenost kanala veze je 93.756%.

Kodirana poruka BADCDBCBCB glasi (razmak između svakog slova):
001 01 000 1 000 001 1 001 1 001

c) Binarni Shannon-Fano kod sa simbolima 0 i 1 (parovi poruka):

CC 0.2025	0.53/0	0.315/00	0.2025/000				
AC 0.1125			0.1125/001				
CA 0.1125		0.225/01	0.1125/010				
CD 0.1125			0.1125/011				
DC 0.1125	0.46/1	0.2375/10	0.1125/100				
AD 0.0625			0.125/101	0.0625/1010			
DA 0.0625				0.0625/1011			
DD 0.0625		0.2225/11	0.125/110	0.0625/1100			
AA 0.0625				0.0625/1101			
BC 0.0225			0.045/1110	0.0225/11100			
CB 0.0225				0.0225/11101			
BD 0.0125			0.0975/111	0.025/11110	0.0125/111100		
DB 0.0125					0.0125/111101		
AB 0.0125				0.0275/11111	0.0125/111110		
BA 0.0125					0.015/111111	0.0125/1111110	
BB 0.0025						0.0025/1111111	

Iz tabele se vidi koji su parovi poruka kodirani kojim kodom, izračunajmo prosječnu dužinu kodne riječi:

$$n_{sr} = 0.2025 \cdot 3 + 4 \cdot 0.1125 \cdot 3 + 4 \cdot 0.0625 \cdot 4 + 2 \cdot 0.0225 \cdot 5 + 3 \cdot 0.0125 \cdot 6 + 0.0125 \cdot 7 + 0.0025 \cdot 7$$

$$n_{sr} = 3.5125$$

Entropija izvora je ovdje faktički entropija sekvenci dužine 2, s obzirom da ne postoji zavisnost unazad. Pored toga, zbog nepostojanja zavisnosti unazad također vrijedi i $H(X^2) = 2 \cdot H(X)$, tako da za protok informacija dobijamo

$$\overline{I(X)} = \frac{2 \cdot H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{0.98761}{\tau}$$

odnosno iskorištenost kanala veze je 98.761%.

Kodirana poruka BADCDBCBCB glasi (razmak između svaka 2 slova):

1111110 100 111101 11101 11101

d) Binarni Huffmanov kod sa simbolima 0 i 1 (parovi poruka):

CC 0.2025	CC 0.2025	CC 0.2025	CC 0.2025	CC 0.2025	CC 0.2025	CC 0.2025	CC 0.2025
AC 0.1125	AC 0.1125	AC 0.1125	AC 0.1125	AC 0.1125	AC 0.1125	AC 0.1125	DD/0
CA 0.1125	CA 0.1125	CA 0.1125	CA 0.1125	CA 0.1125	CA 0.1125	CA 0.1125	AA/1 0.125
CD 0.1125	CD 0.1125	CD 0.1125	CD 0.1125	CD 0.1125	CD 0.1125	CD 0.1125	AC 0.1125
DC 0.1125	DC 0.1125	DC 0.1125	DC 0.1125	DC 0.1125	DC 0.1125	DC 0.1125	CA 0.1125
AD 0.0625	AD 0.0625	AD 0.0625	AD 0.0625	AD 0.0625	AD 0.0625	BA/0000 BB/0001 BD/001	CD 0.1125
DA 0.0625	DA 0.0625	DA 0.0625	DA 0.0625	DA 0.0625	DA 0.0625		DC 0.1125
DD 0.0625	DD 0.0625	DD 0.0625	DD 0.0625	DD 0.0625	DD 0.0625		BA/0000 BB/0001 BD/001 DB/010 0.0975 AB/011 BC/10 CB/11 AD 0.0625 DA 0.0625 DD 0.0625 AA 0.0625
AA 0.0625	AA 0.0625	AA 0.0625	AA 0.0625	AA 0.0625	AA 0.0625	DB/010 0.0975	
BC 0.0225	BC 0.0225	DB/0 0.025	BA/00 BB/01 0.0275	BC/0 0.045	BA/00 BB/001 0.0525	AB/011	
CB 0.0225	CB 0.0225	AB/1		CB/1		BC/10	
BD 0.0125	BA/0 0.015 BB/1	BC 0.0225	BD/1	BA/00	BD/01 DB/10 AB/11	CB/11	
DB 0.0125		CB 0.0225	DB/0 0.025 AB/1	BB/01 0.0275		AD 0.0625	
AB 0.0125	BD 0.0125	BA/0 0.015		BD/1		DA 0.0625	
BA 0.0125	DB 0.0125	BB/1	BC 0.0225	DB/0 0.025	BC/0 0.045	DD 0.0625	AD 0.0625
BB 0.0025	AB 0.0125	BD 0.0125	CB 0.0225	AB/1	CB/1	AA 0.0625	DA 0.0625

CC 0.2025	DC/0	CA/0	DD/00	CC/0	CA/00	CC/00	CC/000
AD/0	BA/10000	CD/1 0.225	AA/01 0.2375	AD/10 0.3275	CD/01	AD/010	AD/0010
DA/1 0.125	BB/10001	DC/0	AC/1	DA/11	DC/10	DA/011	DA/0011
DD/0	BD/1001	BA/10000	CA/0	DD/00	BA/110000	DD/100 0.565	DD/0100
AA/1 0.125	DB/1010 0.21	BB/10001	CD/1 0.225	AA/01 0.2375	BB/110001	AA/101	AA/0101
AC 0.1125	AB/1011	BD/1001	DC/0	AC/1	BD/11001 0.435	AC/11	AC/011
CA 0.1125	BC/110	DB/1010 0.21	BA/10000	CA/0	DB/11010	CA/00	CA/100
CD 0.1125	CB/111	AB/1011	BB/10001	CD/1 0.225	AB/11011	CD/01	CD/101
DC 0.1125	CC 0.2025	BC/110	BD/1001	DC/0	BC/1110	DC/10	DC/110
BA/0000	AD/0	CB/111	DB/1010 0.21	BA/10000	CB/1111	BA/110000	BA/1110000
BB/0001	DA/1 0.125	CC 0.2025	AB/1011	BB/10001	CC/0	BB/110001	BB/1110001
BD/001	DD/0	AD/0	BC/110	BD/1001	AD/10 0.3275	BD/11001 0.435	BD/111001
DB/010 0.0975	AA/1 0.125	DA/1 0.125	CB/111	DB/1010 0.21	DA/11	DB/11010	DB/111010
AB/011	AC 0.1125	DD/0	CC 0.2025	AB/1011	DD/00	AB/11011	AB/111011
BC/10	CA 0.1125	AA/1 0.125	AD/0	BC/110	AA/01 0.2375	BC/1110	BC/11110
CB/11	CD 0.1125	AC 0.1125	DA/1 0.125	CB/111	AC/1	CB/1111	CB/11111

Iz tabele se vidi koji su parovi poruka kodirani kojim kodom, izračunajmo prosječnu dužinu kodne riječi:

$$n_{sr} = 0.2025 \cdot 3 + 4 \cdot 0.0625 \cdot 4 + 4 \cdot 0.1125 \cdot 3 + 7 \cdot 0.0125 + 7 \cdot 0.0025 + 3 \cdot 6 \cdot 0.0125 \\ + 2 \cdot 5 \cdot 0.0225 = 3.5125$$

Entropija izvora je ovdje faktički entropija sekvenci dužine 2, s obzirom da ne postoji zavisnost unazad. Pored toga, zbog nepostojanja zavisnosti unazad također vrijedi i $H(X^2) = 2 \cdot H(X)$, tako da za protok informacija dobijamo

$$\overline{I(X)} = \frac{2 \cdot H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{0.98761}{\tau}$$

odnosno iskorištenost kanala veze je 98.761%.

Kodirana poruka BADCDBCBCB glasi (razmak između svaka 2 slova):

1110000 110 111010 11111 11111

7. Rješenje zadatka

Zadatak 7 [0.8 poena]

Markovljev izvor informacija prvog reda emitira tri različite poruke a, b i c. Ovisno od toga koja je poruka posljednja emitirana, izvor se nalazi u jednom od 3 moguća stanja S_a , S_b i S_c koja redom odgovaraju emitiranim porukama a, b odnosno c. Vjerovatnoće da će izvor emitirati neku od ove 3 poruke ovisno od stanja u kojem se nalazi date su u sljedećoj tablici:

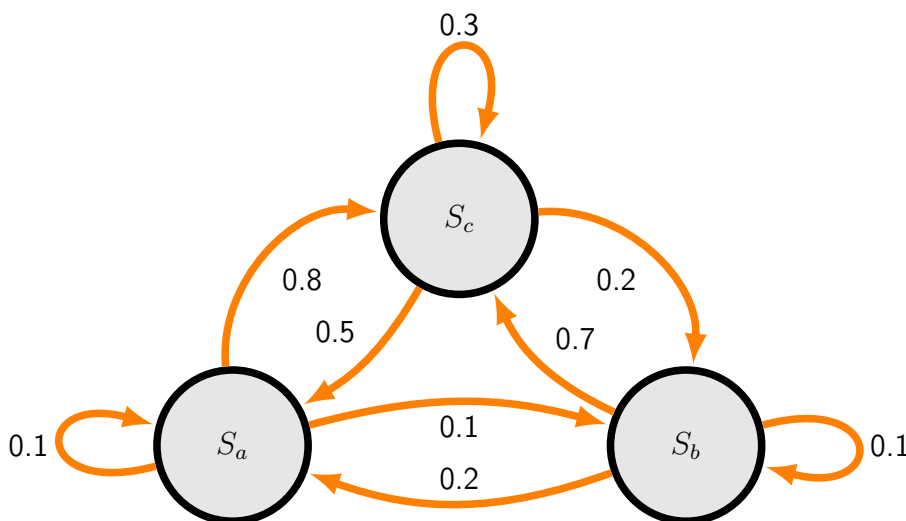
$p(x_j / S_i)$	a	b	c
S_a	0.1	0.1	0.8
S_b	0.2	0.1	0.7
S_c	0.5	0.2	0.3

Za ovaj izvor informacija formirajte binarni Shannon-Fano kod sa simbolima 0 i 1

- posmatrajući izvor kao izvor bez memorije;
- posmatrajući izvor kao izvor bez memorije, ali kodirajući parove poruka umjesto individualnih poruka;
- koristeći posebno kodiranje za svako stanje;
- koristeći posebno kodiranje za svako stanje, ali kodirajući parove poruka umjesto individualnih poruka.

Za sva četiri načina kodiranja, izračunajte protok informacija kroz komunikacioni kanal, procenat iskorištenja kanala veze, te kodirajte sekvencu poruka bccaccabb. U posljednja dva slučaja, pretpostavite da izvor započinje rad u stanju S_a .

* Grafički prikaz konačnog automata koji modelira ovaj izvor dat je na slici ispod:



Na osnovu teoreme o totalnoj vjerovatnoći imamo:

$$P(S_a) = P(S_a)P(a/S_a) + P(S_b)P(a/S_b) + P(S_c)P(a/S_c)$$

$$P(S_b) = P(S_a)P(b/S_a) + P(S_b)P(b/S_b) + P(S_c)P(b/S_c)$$

$$P(S_a) + P(S_b) + P(S_c) = 1$$

Uvrštavanjem vrijednosti iz tabele iznad dobijamo sistem jednačina čija su rješenja:

$$P(S_a) = 0.3245 \quad P(S_b) = 0.1523 \quad P(S_c) = 0.5231$$

Izračunajmo sada entropije svakog od stanja S_a , S_b i S_c čitajući vrijednosti redova iz tabele i koristeći formulu za entropiju:

$$H(S_a) = 0.9219 \quad H(S_b) = 1.1567 \quad H(S_c) = 1.4854$$

Sada možemo izračunati entropiju izvora:

$$H(X/X^\infty) = P(S_a)H(S_a) + P(S_b)H(S_b) + P(S_c)H(S_c) = 1.2523$$

- Pređimo sada na kodiranje ($S_a \rightarrow a$, $S_b \rightarrow b$, $S_c \rightarrow c$):

a) posmatrajući izvor kao izvor bez memorije:

c 0.5231	0.5231/0	
a 0.3245	0.4768/1	0.3245/10
b 0.1523		0.1523/11

Iz tabele se vidi koji su parovi poruka kodirani kojim kodom, izračunajmo prosječnu dužinu kodne riječi:

$$n_{sr} = 0.5231 + 0.3245 \cdot 2 + 0.1523 \cdot 2 = 1.4767$$

samim tim protok je:

$$\overline{I(X)} = \frac{H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{0.8480}{\tau}$$

odnosno iskorištenost je: 84.80%.

Kodirana poruka bccaccabb glasi (razmak između svakog slova):

11 0 0 10 0 0 0 10 11 11

b) posmatrajući izvor kao izvor bez memorije, ali kodirajući parove poruka umjesto individualnih poruka:

cc 0.2736	0.4433/0	0.2736/00			
ac 0.1697		0.1697/01			
ca 0.1697	0.55619/1	0.275/10	0.1697/100		
aa 0.1053			0.1053/101		
bc 0.0796		0.281/11	0.1592/110	0.0796/1100	
cb 0.0796				0.0796/1101	
ab 0.0494			0.121/111	0.0494/1110	
ba 0.0494				0.07259/1111	0.0494/11110
bb 0.02319					0.02319/11111

Iz tabele se vidi koji su parovi poruka kodirani kojim kodom, izračunajmo prosječnu dužinu kodne riječi:

$$n_{sr} = 0.2736 \cdot 2 + 0.1697 \cdot 2 + 0.1697 \cdot 3 + 0.1053 \cdot 3 + 2 \cdot 0.0796 \cdot 4 + 0.0494 \cdot 4 + 0.0494 \cdot 5 + 0.02319 \cdot 5$$

$$n_{sr} = 2.90895$$

samim tim protok je:

$$\overline{I(X)} = \frac{2 \cdot H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{0.8609}{\tau}$$

odnosno iskorištenost je: 86.09%.

Kodirana poruka bccacccabb glasi (razmak između svaka 2 slova):

1100 100 00 100 11111

c) koristeći posebno kodiranje za svako stanje:

S_a

c 0.8	0.8/0	
b 0.1	0.2/1	0.1/10
a 0.1		0.1/11

$$n_{sr_a} = 0.8 + 2 \cdot 0.1 \cdot 2 = 1.2$$

S_b

c 0.7	0.7/0	
b 0.2	0.3/1	0.2/10
a 0.1		0.1/11

$$n_{sr_b} = 0.7 + 0.2 \cdot 2 + 0.1 \cdot 2 = 1.3$$

S_c

a 0.5	0.5/0	
c 0.3	0.5/1	0.3/10
b 0.2		0.2/11

$$n_{sr_c} = 0.5 + 0.3 \cdot 2 + 0.2 \cdot 2 = 1.5$$

$$n_{sr} = P(S_a) \cdot n_{sr_a} + P(S_b) \cdot n_{sr_b} + P(S_c) \cdot n_{sr_c} = 0.3245 \cdot 1.2 + 0.1523 \cdot 1.3 + 0.5231 \cdot 1.5$$

$$n_{sr} = 1.37204$$

Pošto izvor započinje rad u stanju S_a kodirana poruka bccacccabb glasi (razmak po slovu):

10 0 10 0 0 10 10 0 10 10

Protok informacija je:

$$\overline{I(X)} = \frac{H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{0.9127}{\tau}$$

Iskorištenost kanala veze je približno 91.27%.

d) koristeći posebno kodiranje za svako stanje, ali kodirajući parove poruka umjesto individualnih poruka:

S_a

cc 0.64	0.64/0					
ac 0.08	0.36/1	0.16/10	0.08/100			
ca 0.08			0.08/101			
bc 0.08		0.2/11	0.08/110			
cb 0.08			0.12/111	0.08/1110		
ab 0.01				0.04/1111	0.02/11110	0.01/111100
ba 0.01					0.02/11111	0.01/111101
bb 0.01						0.01/111110
aa 0.01						0.01/111111

$$n_{sr_a} = 0.64 + 3 \cdot 0.08 \cdot 3 + 0.08 \cdot 4 + 4 \cdot 0.01 \cdot 6 = 1.92$$

S_b

cc 0.49	0.49/0					
ac 0.14	0.51/1	0.28/10	0.14/100			
ca 0.14			0.14/101			
bc 0.07		0.23/11	0.14/110	0.07/1100		
cb 0.07				0.07/1101		
aa 0.04			0.09/111	0.04/1110		
ab 0.02				0.05/1111	0.02/11110	
ba 0.02					0.03/11111	0.02/111110
bb 0.01						0.01/111111

$$n_{sr_b} = 0.49 + 2 \cdot 0.14 \cdot 3 + 2 \cdot 0.07 \cdot 4 + 0.04 \cdot 4 + 0.02 \cdot 5 + 0.02 \cdot 6 + 0.01 \cdot 6 = 2.33$$

S_c

aa 0.25	0.55/0	0.25/00			
ac 0.15		0.3/01	0.15/010		
ca 0.15			0.15/011		
ab 0.1	0.45/1	0.2/10	0.1/100		
ba 0.1			0.1/101		
cc 0.09		0.25/11	0.15/110	0.09/1100	
bc 0.06				0.06/1101	
cb 0.06			0.1/111	0.06/1110	
bb 0.04				0.04/1111	

$$n_{sr_c} = 0.25 \cdot 2 + 2 \cdot 0.15 \cdot 3 + 2 \cdot 0.1 \cdot 3 + 0.09 \cdot 4 + 2 \cdot 0.06 \cdot 4 + 0.04 \cdot 4 = 3$$

Izračunajmo n_{sr} :

$$n_{sr} = P(S_a) \cdot n_{sr_a} + P(S_b) \cdot n_{sr_b} + P(S_c) \cdot n_{sr_c} = 0.3245 \cdot 1.92 + 0.1523 \cdot 2.33 + 0.5231 \cdot 3$$

$$n_{sr} = 2.547199$$

Protok informacija je:

$$\overline{I(X)} = \frac{2 \cdot H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{0.98327}{\tau}$$

Iskorištenost kanala veze je približno 98.327%.

Pošto izvor započinje rad u stanju S_a kodirana poruka bccacccabb glasi (razmak po svakom markovljevom lancu):

110 011 0 011 111110

8. Rješenje zadatka

Zadatak 8 [0.25 poena]

Neki binarni kanal veze sa šumom prenosi dva simbola 0 i 1, pri čemu su vjerovatnoće greške nule i jedinice 0.1 i 0.05 respektivno. Odredite količinu prenesene informacije kroz ovaj kanal ukoliko vjerovatnoća pojave nule na ulazu u kanal iznosi 0.8, te odredite njegov kapacitet.

* Ulazne simbole označimo sa $y_1 = 0$ i $y_2 = 1$, a izlazne simbole sa $z_1 = 0$ i $z_2 = 1$. Iz postavke zadatka imamo:

$$p(z_1/y_1) = 0.9 \quad p(z_2/y_1) = 0.1 \quad p(z_1/y_2) = 0.05 \quad p(z_2/y_2) = 0.95$$

odnosno:

$$p(y_1) = 0.8 \quad p(y_2) = 0.2$$

Na osnovu teoreme o totalnoj vjerovatnoći, za vjerovatnoće simbola na izlazu iz kanala imamo:

$$p(z_1) = p(y_1) \cdot p(z_1/y_1) + p(y_2) \cdot p(z_1/y_2) = 0.8 \cdot 0.9 + 0.2 \cdot 0.05 = 0.73$$

$$p(z_2) = 1 - p(z_1) = 1 - 0.73 = 0.27$$

Za entropije ulaznih i izlaznih simbola imamo:

$$H(Y) = -(p(y_1) \cdot \log_2 p(y_1) + p(y_2) \cdot \log_2 p(y_2)) = 0.7219$$

$$H(Z) = -(p(z_1) \cdot \log_2 p(z_1) + p(z_2) \cdot \log_2 p(z_2)) = 0.8414$$

Za računanje uvjetne entropije $H(Z/Y)$, koja nam treba za računanje prenesene količine informacija, prvo ćemo odrediti "djelimično" uvjetne entropije $H(Z/y_1)$ i $H(Z/y_2)$:

$$H(Z/y_1) = -(p(z_1/y_1) \cdot \log_2(p(z_1/y_1))) + p(z_2/y_1) \cdot \log_2(p(z_2/y_1)) = 0.4689$$

$$H(Z/y_2) = -(p(z_1/y_2) \cdot \log_2(p(z_1/y_2))) + p(z_2/y_2) \cdot \log_2(p(z_2/y_2)) = 0.2863$$

Oдавде za uvjetnu entropiju $H(Z/Y)$ dobijamo:

$$H(Z/Y) = p(y_1) \cdot H(Z/y_1) + p(y_2) \cdot H(Z/y_2) = 0.8 \cdot 0.4689 + 0.2 \cdot 0.2863 = 0.43238$$

Prenesena količina informacije kroz kanal je data sa:

$$I(Y, Z) = H(Z) - H(Z/Y) = 0.8414 - 0.43238 = 0.40902$$

Kapacitet kanala veze je dat relacijom:

$$C_c = \frac{1}{\tau} \cdot \left[\log_2 \left(2^{-\frac{H_0}{\gamma}} + 2^{-\frac{H_1}{\gamma}} \right) + \frac{H_0 \cdot \beta + H_1 \cdot \alpha}{\gamma} \right]$$

gdje su:

$$\alpha = 0.1 \quad \beta = 0.05 \quad H_0 = 0.4689 \quad H_1 = 0.2863 \quad \gamma = 1 - \alpha - \beta = 0.85$$

Uvrštavanjem navedenih vrijednosti u formulu iznad, dobijamo da je kapacitet:

$$C_c = \frac{0.62102}{\tau}$$

9. Rješenje zadatka

Zadatak 9 [0.4 poena]

Izvor informacija bez memorije emitira dvije poruke a i b, pri čemu vjerovatnoća emitiranja poruke a iznosi $p(a) = 0.8$. Ove poruke se zatim kodiraju, i prenose kroz binarni kanal veze sa šumom koji koristi dva simbola 0 i 1, pri čemu su vjerovatnoće greške nule i jedinice 0.25 i 0.05 respektivno.

Odredite količinu prenesene informacije kroz komunikacioni kanal, brzinu prenosa informacija kroz komunikacioni kanal, procenat iskorištenja kanala veze i vjerovatnoću greške u prenosu ukoliko se koristi:

- Prosto kodiranje a \rightarrow 0 i b \rightarrow 1 uz dekodiranje zasnovano na klasifikaciji $S_a = \{0\}$ i $S_b = \{1\}$;
- Zaštitno kodiranje a \rightarrow 000 i b \rightarrow 111 uz dekodiranje zasnovano na klasifikaciji $S_a = \{000, 001, 010, 100\}$ i $S_b = \{011, 101, 110, 111\}$.

* Ulazne simbole označimo sa $y_1 = 0$ i $y_2 = 1$, a izlazne simbole sa $z_1 = 0$ i $z_2 = 1$. Iz postavke zadatka imamo:

$$p(z_1/y_1) = 0.75 \quad p(z_2/y_1) = 0.25 \quad p(z_1/y_2) = 0.05 \quad p(z_2/y_2) = 0.95$$

odnosno:

$$p(a) = 0.8 \quad p(b) = 0.2$$

Neka su poruke koje prima krajnji korisnik w_1 i w_2 .

a) Kako se kodiranje vrši prostim ravnomjernim kodom a $\rightarrow 0$ i b $\rightarrow 1$ (tj. a $\rightarrow y_1$ i b $\rightarrow y_2$) uz dekodiranje zasnovano na klasifikaciji $S_a = \{0\}$ i $S_b = \{1\}$ (tj. $S_a = \{z_1\}$ i $S_b = \{z_2\}$).

Kako se w_1 dekodira samo ako je primljeno $z_1 = 0$, a w_2 samo ako je primljeno $z_2 = 1$ slijedi:

$$\begin{aligned} p(w_1/a) &= p(z_1/y_1) = 0.75 \\ p(w_2/a) &= 1 - p(w_1/a) = 0.25 \\ p(w_1/b) &= p(z_1/y_2) = 0.05 \\ p(w_2/b) &= 1 - p(w_1/b) = 0.95 \end{aligned}$$

Dalje imamo:

$$p(w_1) = p(a) \cdot p(w_1/a) + p(b) \cdot p(w_1/b) = 0.8 \cdot 0.75 + 0.2 \cdot 0.05 = 0.61$$

$$p(w_2) = 1 - p(w_1) = 0.39$$

Sada možemo izračunati naredno:

$$\begin{aligned} H(W) &= -(p(w_1) \cdot \log_2 p(w_1) + p(w_2) \cdot \log_2 p(w_2)) = 0.9648 \\ H(W/a) &= -(p(w_1/a) \cdot \log_2 p(w_1/a) + p(w_2/a) \cdot \log_2 p(w_2/a)) = 0.811278 \\ H(W/b) &= -(p(w_1/b) \cdot \log_2 p(w_1/b) + p(w_2/b) \cdot \log_2 p(w_2/b)) = 0.28634 \\ H(W/X) &= p(a) \cdot H(W/a) + p(b) \cdot H(W/b) = 0.7063 \\ I(X, W) &= H(W) - H(W/X) = 0.9648 - 0.7063 = 0.2585 \end{aligned}$$

Odredimo brzinu prenosa informacija kroz komunikacioni kanal:

$$\overline{I(X, W)} = \frac{I(X, W)}{\tau} = \frac{0.2585}{\tau}$$

Procenat iskorištenosti kanala veze je 60.911% (iskorištenost u prenosu / $C_c - C_c$ dobijemo po ogromnoj formuli iz 8 zadatka). Izračunajmo još vjerovatnoću greške u prenosu:

$$p_e = 1 - p(a) \cdot p(w_1/a) - p(b) \cdot p(w_2/b) = 1 - 0.8 \cdot 0.75 - 0.2 \cdot 0.95 = 0.21$$

b) Kodiranje se vrši zaštitnim kodom a \rightarrow 000 i b \rightarrow 111 (tj. a \rightarrow $y_1y_1y_1$ i b \rightarrow $y_2y_2y_2$) uz dekodiranje zasnovano na klasifikaciji $S_a = \{000, 001, 010, 100\}$ i $S_b = \{011, 101, 110, 111\}$. (tj. $S_a = \{z_1z_1z_1, z_1z_1z_2, z_1z_2z_1, z_2z_1z_1\}$ i $S_b = \{z_1z_2z_2, z_2z_1z_2, z_2z_2z_1, z_2z_2z_2\}$). Uz ovakvo kodiranje i dekodiranje imamo:

$$\begin{aligned} p(w_1/a) &= p(z_1z_1z_1/a) + p(z_1z_1z_2/a) + p(z_1z_2z_1/a) + p(z_2z_1z_1/a) = \\ &= p(z_1z_1z_1/y_1y_1y_1) + p(z_1z_1z_2/y_1y_1y_1) + p(z_1z_2z_1/y_1y_1y_1) + p(z_2z_1z_1/y_1y_1y_1) = \\ &= p(z_1/y_1)^3 + 3 \cdot (z_1/y_1)^2 \cdot p(z_2/y_1) = 0.75^3 + 3 \cdot 0.75^2 \cdot 0.25 = 0.84375 \\ p(w_2/a) &= 1 - p(w_1/a) = 1 - 0.84375 = 0.15625 \end{aligned}$$

Na isti način dobijamo da je

$$\begin{aligned} p(w_1/b) &= p(z_1/y_2)^3 + 3 \cdot (z_1/y_2)^2 \cdot p(z_2/y_2) = 0.05^3 + 3 \cdot 0.05^2 \cdot 0.95 = 0.00725 \\ p(w_2/b) &= 1 - 0.00725 = 0.99275 \end{aligned}$$

Dalje imamo:

$$\begin{aligned} p(w_1) &= p(a) \cdot p(w_1/a) + p(b) \cdot p(w_1/b) = 0.8 \cdot 0.84375 + 0.2 \cdot 0.00725 = 0.67645 \\ p(w_2) &= 1 - p(w_1) = 0.32355 \end{aligned}$$

Sada možemo izračunati naredno:

$$\begin{aligned} H(W) &= -(p(w_1) \cdot \log_2 p(w_1) + p(w_2) \cdot \log_2 p(w_2)) = 0.9082 \\ H(W/a) &= -(p(w_1/a) \cdot \log_2 p(w_1/a) + p(w_2/a) \cdot \log_2 p(w_2/a)) = 0.62526 \\ H(W/b) &= -(p(w_1/b) \cdot \log_2 p(w_1/b) + p(w_2/b) \cdot \log_2 p(w_2/b)) = 0.06139 \\ H(W/X) &= p(a) \cdot H(W/a) + p(b) \cdot H(W/b) = 0.512486 \\ I(X, W) &= H(W) - H(W/X) = 0.9082 - 0.512486 = 0.395714 \end{aligned}$$

Odredimo brzinu prenosa informacija kroz komunikacioni kanal:

$$\overline{I(X, W)} = \frac{I(X, W)}{3 \cdot \tau} = \frac{0.395714}{3 \cdot \tau} = \frac{0.1319}{\tau}$$

Procenat iskorištenosti kanala veze je 20.88% (iskorištenost u prenosu / C_c - C_c dobijemo po ogromnoj formuli iz 8 zadatka). Izračunajmo još vjerovatnoću greške u prenosu:

$$p_e = 1 - p(a) \cdot p(w_1/a) - p(b) \cdot p(w_2/b) = 1 - 0.8 \cdot 0.84375 - 0.2 \cdot 0.99275 = 0.12645$$