Zadaća 3

iz predmeta Diskretna Matematika

Prezime i ime: Mašović Haris

Br. indexa: 17993

Demonstrator: Rijad Muminović

Grupa: RI - 4

Zadatak	Bodovi
1	
2	
3	
4	
5	
6	
7	
8	
9	

Elektrotehnički fakultet Sarajevo, Decembar 2017g.

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 A1, A2 i A3. 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	0.2	0.35	0.05	0.1	0.3
A_2	0.45	0.15	0.1	0.15	0.15
A_3	0.45	0.1	0.3	0.05	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.

* S obzirom da imamo vjerovatnoće događaja X možemo odmah izračunati entropiju H(X):

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

odnosno:

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

Da bi izračunali entropiju H(Y) potrebne su nam sve vjerovatnoće P(Y_j), j $\in$ {1,..,5}. Za računanje tih vjerovantoća koristit ćemo formulu

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

odnosno to možemo uraditi sumirajući kolone preko tabele:

p($B_j A_i$)	B_1	B_2	B_3	B_4	B_5
A_1	0.04	0.07	0.01	0.02	0.06
A_2	0.27	0.09	0.06	0.09	0.09
A_3	0.09	0.02	0.06	0.01	0.02

Sada možemo izračunati vjerovatnoće P(Y_j); j = 1..5 i to:

$$P(B_1) = 0.04 + 0.27 + 0.09 = 0.4 \quad P(B_2) = 0.07 + 0.09 + 0.02 = 0.18$$

$$P(B_3) = 0.01 + 0.06 + 0.06 = 0.13 \quad P(B_4) = 0.02 + 0.09 + 0.01 = 0.12$$

$$P(B_5) = 0.06 + 0.09 + 0.02 = 0.17$$

Sada možemo izračunati H(Y):

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

odnosno:

$$H(Y) = -(0.4 \cdot log_2\ 0.4 + 0.18 \cdot log_2\ 0.18 + 0.13 \cdot log_2\ 0.13 + 0.12 \cdot log_2\ 0.12 + 0.17 \cdot log_2\ 0.17)$$

odnosno:

$$H(Y) = 2.15838$$

Izračunati ćemo H(X,Y) preko gornje tabele tako da posmatramo svaki broj unutar tabele:

$$H(X,Y) = -\sum_{i=1}^{3}\sum_{j=1}^{5} P(B_jA_i) \cdot log_2\ (P(B_jA_i)) = 3.417059$$

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

$$H(X/Y) = H(X,Y) - H(Y) = 3.417059 - 2.15838 = 1.258679$$

$$H(Y/X) = H(X,Y) - H(X) = 3.417059 - 1.37095 = 2.046109$$

$$I(X,Y) = H(X) - H(X/Y) = 1.37095 - 1.258679 = 0.112271$$

2. Rješenje zadatka

Zadatak 2 [0.25 poena]

Na nekom fakultetu, troškove studija za 37% studenata plaća država, dok su ostali studenti samofinansirajući. Među studentima koji se školuju o trošku države, 54% 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 44% 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 događaja "troškove studija studenta plaća država" i "student je samofinansirajući" možemo označiti sa X, pri čemu će prvi događaj biti X_1 a drugi događaj X_2. Na osnovu postavke zadatka, poznato je da za ove događaje vrijedi da je

$$p(X_1) = 0.37 \quad p(X_2) = 0.63$$

Pored ovog skupa događaja, može se formirati i skup događaja Y1, Y2 označen sa Y, čiji elementi respektivno označavaju događaje "student stanuje u studentskom domu" i "student ne stanuje u studentskom domu". Zadane su uslovne vjerovatnoće

$$p(Y_1/X_1) = 0.54 \quad p(Y_1/X_2) = 0.39$$

Pošto je neophodno da vrijedi da svaki student čije troškove studija plaća država ili stanuje ili ne stanuje u studentskom domu, te analogno za studente koji su samofinansirajući, vrijede sljedeće relacije:

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

na osnovu čega se mogu izračunati tražene vjerovatnoće:

$$p(Y_2/X_1) = 1 - p(Y_1/X_1) = 0.46 \quad p(Y_2/X_2) = 1 - p(Y_1/X_2) = 0.61$$

Dalje, potrebno je uvesti i skup događaja Z sa događajima $\{Z_1, Z_2\}$ koji respektivno predstavljaju tvrdnje „student posjeduje iskaznicu za subvencionirani javni prevoz“ i „student ne posjeduje iskaznicu za subvencionirani javni prevoz“. S obzirom na činjenicu da svi studenti koji stanuju u studentskom domu također posjeduju iskaznicu za subvencionirani javni prevoz, dolazi se do zaključka da najmanje 54% studenata čije troškove studija plaća država te 39% samofinansirajućih studenata posjeduje iskaznicu. Također, s obzirom na to da i među studentima koji ne stanuju u domu iskaznicu ima 44% studenata sa državnog budžeta, vjerovatnoća da student posjeduje iskaznicu uz uslov da mu troškove studija plaća država računa se kao: 0.54 + 0.44 · (1 - 0.54) = 0.7424. Dobijena vjerovatnoća po već uvedenim oznakama predstavlja vjerovatnoću p(Z_1/X_1). Jasno je da svaki student čije troškove studija plaća država ili ima ili nema iskaznicu, pa je vjerovatnoća p(Z_2/X_1) = 1 - 0.7424 = 0.2576.

Na analogan način računa se vjerovatnoća da student posjeduje iskaznicu uz uslov da je samofinansirajući, te ona iznosi 0.39 + 0.45 · (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. Za četiri dobijene vrijednosti najpreglednije je formirati tabelu:

p(Z_j/X_i)	Z_1	Z_2
X_1	0.7424	0.2576
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 može se računati uz pomoć formule:

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

Izračunajmo H(X) kao:

$$H(X) = -(p(X_1) \cdot log_2\, p(X_1) + p(X_2) \cdot log_2\, p(X_2)) = -(0.37 \cdot log_2\, 0.37 + 0.63 \cdot log_2\, 0.63) = 0.950672$$

Izračunajmo H(Z) sada:

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

Vjerovatnoće p(Z_1) i p(Z_2) nisu poznate, ali mogu se naći i to:

$$p(Z_1) = p(X_1) \cdot p(Z_1/X_1) + p(X_2) \cdot p(Z_1/X_2) = 0.37 \cdot 0.7424 + 0.63 \cdot 0.6645 = 0.693323$$

$$p(Z_2) = p(X_1) \cdot p(Z_2/X_1) + p(X_2) \cdot p(Z_2/X_2) = 0.37 \cdot 0.2576 + 0.63 \cdot 0.3355 = 0.306677$$

Entropiju H(Z) možemo sad izračunati:

$$H(Z) = -(0.693323 \cdot log_2\, 0.693323 + 0.306677 \cdot log_2\, 0.306677) = 0.889300$$

Izračunajmo H(X,Z):

$$H(X,Z) = -(p(X_1Z_1) \cdot log_2\, p(X_1Z_1) + p(X_2Z_1) \cdot log_2\, p(X_2Z_1) +$$

$$p(X_1Z_2) \cdot log_2\, p(X_1Z_2) + p(X_2Z_2) \cdot log_2\, p(X_2Z_2))$$

p(X_i Z_j) smo već faktički izračunali kod p(Z_i) samim tim možemo samo uvrstiti pa imamo:

$$H(X,Z) = -(0.274688 \cdot log_2\, 0.274688 + 0.418635 \cdot log_2\, 0.418635 +$$

$$0.095312 \cdot log_2\, 0.095312 + 0.211365 \cdot log_2\, 0.211365) = 1.83510$$

Na kraju I(X,Z) imamo da je:

$$I(X,Z) = H(X) + H(Z) - H(X,Z) = 0.950672 + 0.889300 - 1.83510 = 0.004872$$

3. Rješenje zadatka

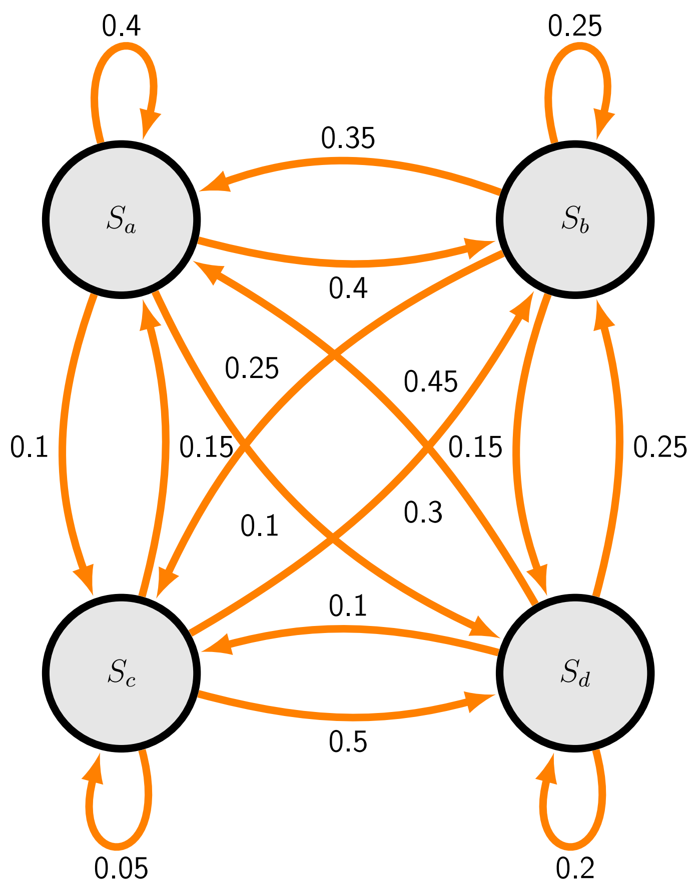
Zadatak 3 [0.35 poena]

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 Sa, Sb, Sc i Sd 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	0.4	0.1	0.1
S_b	0.35	0.25	0.25	0.15
S_c	0.15	0.3	0.05	0.5
S_d	0.45	0.25	0.1	0.2

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

* Kako je red izvora r = 1, 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 imamo:

$$P(S_a) = P(S_a) \cdot P(a/S_a) + P(S_b) \cdot P(a/S_b) + P(S_c) \cdot P(a/S_c) + P(S_d) \cdot P(a/S_d)$$

Na osnovu toga, sa svako stanje imamo:

$$P(S_a) = 0.4 \cdot P(S_a) + 0.35 \cdot P(S_b) + 0.15 \cdot P(S_c) + 0.45 \cdot P(S_d)$$

$$P(S_b) = 0.4 \cdot P(S_a) + 0.25 \cdot P(S_b) + 0.3 \cdot P(S_c) + 0.25 \cdot P(S_d)$$

$$P(S_c) = 0.1 \cdot P(S_a) + 0.25 \cdot P(S_b) + 0.05 \cdot P(S_c) + 0.1 \cdot P(S_d)$$

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

Rješenja našeg sistema su respektivno:

$$P(S_a) = 0.359069 \quad P(S_b) = 0.3108426 \quad P(S_c) = 0.1396442 \quad P(S_d) = 0.1904442$$

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

$$H(S_a) = -(P(a/S_a) \cdot log_2\ P(a/S_a) + P(b/S_a) \cdot log_2\ P(b/S_a) +$$

$$P(c/S_a) \cdot log_2\ P(c/S_a) + P(d/S_a) \cdot log_2\ P(d/S_a))$$

odnosno kao finalna entropija svakog stanja:

$$H(S_a) = 1.72193 \quad H(S_b) = 1.94065 \quad H(S_c) = 1.64773 \quad H(S_d) = 1.81498$$

Za entropiju izraza imamo:

$$H(X/X^\infty) = P(S_a)H(S_a) + P(S_b)H(S_b) + P(S_c)H(S_c) + P(S_d)H(S_d) = 1.797276726642$$

Redudansu izvora računamo kao:

$$R = \frac{log_2\ 4 - H(X/X^\infty)}{log_2\ 4} = \frac{2 - 1.797276726642}{2} = 0.101361636679$$

Vjerovatnoću sekvence ***bdcaa*** možemo izračunati kao

$$P(bdcaa) = P(b) \cdot P(d/b) \cdot P(c/d) \cdot P(a/c) \cdot P(a/a) = 0.3108426 \cdot 0.15 \cdot 0.1 \cdot 0.15 \cdot 0.4 = 0.00027975834$$

odnosno

$$P(bdcaa) = 0.027975834\%$$

Za računanje entropije sekvenci dužine 5 imamo relaciju:

$$H(X^5) = H(X^1) + 4 \cdot H(X/X^\infty)$$

Treba nam još entropija H(X^1), koja iznosi:

$$H(X^1) = -(P(a) \cdot log_2\ P(a) + P(b) \cdot log_2\ P(b) + P(c) \cdot log_2\ P(c) + P(d) \cdot log_2\ P(d)) = 1.90685$$

Izračunajmo sada entropiju sekvenci 5:

$$H(X^5) = H(X^1) + 4 \cdot H(X/X^\infty) = 1.90685 + 4 \cdot 1.797276726642 = 9.095956906568$$

4. Rješenje zadatka

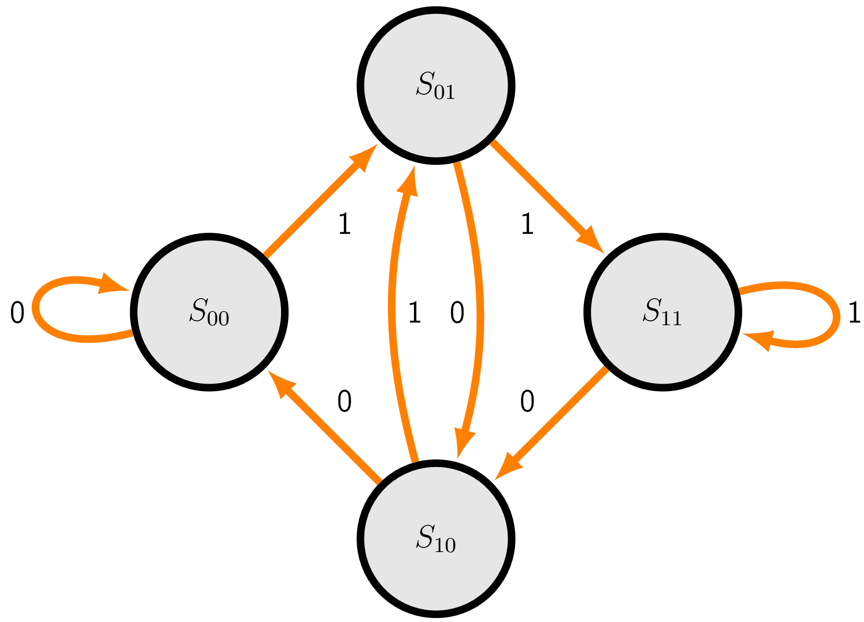
Zadatak 4 [0.4 poena]

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:

$$p(0/S_{00}) = 0.7 \quad p(0/S_{01}) = 0.6 \quad p(0/S_{10}) = 0.1 \quad p(0/S_{11}) = 0.1$$

Odredite entropiju i redudansu ovog izvora, zatim entropiju sekvenci dužine 6 te vjerovatnoću pojave sekvence 00110.

* 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:

$$p(1/S_{00}) = 0.3 \quad p(1/S_{01}) = 0.4 \quad p(1/S_{10}) = 0.9 \quad p(1/S_{11}) = 0.9$$

Izračunajmo entropiju stanja:

$$H(S_{00}) = -(0.7 \cdot log_2\ 0.7 + 0.3 \cdot log_2\ 0.3) = 0.881291$$
$$H(S_{01}) = -(0.6 \cdot log_2\ 0.6 + 0.4 \cdot log_2\ 0.4) = 0.970951$$
$$H(S_{10}) = -(0.1 \cdot log_2\ 0.1 + 0.9 \cdot log_2\ 0.9) = 0.468996$$
$$H(S_{11}) = -(0.1 \cdot log_2\ 0.1 + 0.9 \cdot log_2\ 0.9) = 0.468996$$

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

$$P(S_{00}) = 0.7 \cdot P(S_{00}) + 0.1 \cdot P(S_{10})$$
$$P(S_{01}) = 0.3 \cdot P(S_{00}) + 0.9 \cdot P(S_{10})$$
$$P(S_{11}) = 0.4 \cdot P(S_{01}) + 0.9 \cdot P(S_{11})$$
$$P(S_{00}) + P(S_{01}) + P(S_{10}) + P(S_{11}) = 1$$

Respektivno, rješenja su:

$$P(S_{00}) = 0.05263158 \quad P(S_{01}) = 0.1578947 \quad P(S_{10}) = 0.1578947 \quad P(S_{11}) = 0.6315789$$

Izračunajmo sada entropiju izvora

$$H(X/X^{\infty}) = P(S_{00})H(S_{00}) + P(S_{01})H(S_{01}) + P(S_{10})H(S_{10}) + P(S_{11})H(S_{11}) = 0.56995171513508$$

Redudansa izvora je:

$$R = \frac{log_2 2 - 0.56995171513508}{1} = 0.43004828486492$$

Odredimo entropiju sekvenci dužine 6:

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

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

$$H(X^2) = 1.483226$$

odnosno entropija sekvenci 6 je:

$$H(X^6) = H(X^2) + 4 \cdot H(X/X^\infty) = 1.483226 + 4 \cdot 0.56995171513508 = 3.76303286054032$$

Izračunajmo još i vjerovatnoću sekvence 00110:

$$p(00110) = p(00) \cdot p(1/00) \cdot p(1/01) \cdot p(0/11) = 0.05263158 \cdot 0.3 \cdot 0.4 \cdot 0.1 = 0.00063157896$$

odnosno

$$p(00110) = 0.063157896\%$$

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 521 koju je emitirao ovaj izvor, uočena je sljedeća učestalost pojavljivanja pojedinih poruka:

Poruka:	A	B	C	D	E	F	G	H	I	J
Učestalost:	46	40	62	79	40	30	69	31	29	95

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 DIHFAAHE.

* 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.

95 J	243/0	**95/00**		
79 D		148/01	**79/010**	
69 G			**69/011**	
62 C	278/1	130/10	**62/100**	
46 A			68/101	**46/1010**
40 B				**40/1011**
40 E		148/11	71/110	**40/1100**
31 H				**31/1101**
30 F			59/111	**30/1110**
29 I				**29/1111**

Imamo kodirane sekvence:

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

Izračunajmo H(X/X^∞) pri čemu sad moramo uzeti vjerovatnoće u obzir,
odnosno učestalost / 521:

$$H(X/X^\infty) = -\frac{1}{521}(95 \cdot log_2 \frac{95}{521} + 79 \cdot log_2 \frac{79}{521} + 69 \cdot log_2 \frac{69}{521} + 62 \cdot log_2 \frac{62}{521} + 46 \cdot log_2 \frac{46}{521} +$$

$$40 \cdot log_2 \frac{40}{521} + 40 \cdot log_2 \frac{40}{521} + 31 \cdot log_2 \frac{31}{521} + 30 \cdot log_2 \frac{30}{521} + 29 \cdot log_2 \frac{29}{521})$$

odnosno:

$$H(X/X^\infty) = 3.20116706$$

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:

$$n_{sr} = \frac{1}{521}(95 \cdot 2 + 79 \cdot 3 + 69 \cdot 3 + 62 \cdot 3 + 46 \cdot 4 + 40 \cdot 4 + 40 \cdot 4 + 31 \cdot 4 + 30 \cdot 4 + 29 \cdot 4) = 3.232245$$

pa je protok informacija:

$$\overline{I(X)} = \frac{H(X/X^\infty)}{n_{sr} \cdot \tau} = \frac{0.99038}{\tau}$$

odakle slijedi da je iskorištenost kanala veze približno 99.038%

Kodirana poruka DIHFAAHE glasi (razmak između svakog slova):
010 1111 1101 1110 1010 1010 1101 1100

b) Binarni Huffmanov kod sa 0 i 1:

95 J	95 J	95 J	95 J	C/0 121	E/00 140	A/00 165	C/00 216	A/000	**A/0000**
79 D	79 D	69 D	A/0 86	F/10	H/01	B/01	F/010	B/001	**B/0001**
69 G	69 G	E/0 71	B/1	I/11	G/1	D/1	I/011	D/01	**D/001**
62 C	62 C	H/1	79 D	95 J	C/0 121	E/00 140	J/1	E/100	**E/0100**
46 A	F/0 59	69 G	E/0 71	A/0 86	F/10	H/01	A/00 165	H/101	**H/0101**
40 B	I/1	62 C	H/1	B/1	I/11	G/1	B/01	G/11	**G/011**
40 E	46 A	F/0 59	69 G	79 D	95 J	C/0 121	D/1	C/00	**C/100**
31 H	40 B	I/1	62 C	E/0 71	A/0 86	F/10	E/00 146	F/010	**F/1010**
30 F	40 E	46 A	F/0 59	H/1	B/1	I/11	H/01	I/011	**I/1011**
29 I	31 H	40 B	I/1	69 G	79 D	95 J	G/1	J/1	**J/11**

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

$$n_{sr} = \frac{1}{521}(4 \cdot 46 + 4 \cdot 40 + 3 \cdot 79 + 4 \cdot 40 + 4 \cdot 31 + 3 \cdot 69 + 3 \cdot 62 + 4 \cdot 30 + 4 \cdot 29 + 2 \cdot 95) = 3.232245$$

pa je na osnovu toga protok informacija (pošto je n_{sr} i entropija izvora ista) isti kao u prethodnom zadatku, samim tim je i iskorištenost kanala veze ista.

Kodirana poruka DIHFAAHE glasi (razmak između svakog slova):
001 1011 0101 1010 0000 0000 0101 0100

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

95 J	95 J	B/0 111	C/0 167	J/0 243	**J/00 521**
79 D	79 D	E/1	F/10	D/1	**D/01**
69 G	69 G	H/2	I/11	G/2	**G/02**
62 C	62 C	95 J	A/2	C/0 167	**C/10**
46 A	F/0 59	79 D	B/0 111	F/10	**F/110**
40 B	I/1	69 G	E/1	I/11	**I/111**
40 E	46 A	62 C	H/2	A/2	**A/12**
31 H	40 B	F/0 59	95 J	B/0 111	**B/20**
30 F	40 E	I/1	79 D	E/1	**E/21**
29 I	31 H	46 A	69 G	H/2	**H/22**

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

$$n_{sr} = \frac{1}{521}(2 \cdot 95 + 2 \cdot 79 + 2 \cdot 69 + 2 \cdot 62 + 3 \cdot 30 + 3 \cdot 29 + 2 \cdot 46 + 2 \cdot 40 + 2 \cdot 40 + 2 \cdot 31) = 2.113243$$

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

$$\overline{I(X)} = \frac{H(X/X^{\infty})}{n_{sr} \cdot \tau} = \frac{1.5148125}{\tau}$$

Kako je kapacitet kanala veze $C_c = \frac{log_2 3}{\tau} = \frac{1.5850}{\tau}$ iskorištenost kanala veze je $\frac{1.5148125}{1.5850} = 0.955717 = 95.5717\%$

Kodirana poruka DIHFAAHE glasi (razmak između svakog slova):
01 111 22 110 12 12 22 21

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:

p(A) = 0.15
p(B) = 0.4
p(C) = 0.1
p(D) = 0.35

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

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

B 0.4	**0.4/0**		
D 0.35	0.6/1	**0.35/10**	
A 0.15		0.25/11	**0.15/110**
C 0.1			**0.1/111**

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

$$H(X/X^{\infty}) = -(0.4 \cdot log_2\ 0.4 + 0.35 \cdot log_2\ 0.35 + 0.15 \cdot log_2\ 0.15 + 0.1 \cdot log_2\ 0.1) = 1.80161$$

$$n_{sr} = 0.4 \cdot 1 + 0.35 \cdot 2 + 0.15 \cdot 3 + 0.1 \cdot 3 = 1.85$$

pa je protok kanala veze:

$$\overline{I(X)} = \frac{H(X/X^{\infty})}{n_{sr} \cdot \tau} = \frac{0.9738}{\tau}$$

odnosno iskorištenost kanala veze je 97.38%.

Kodirana poruka DACDACBCDB glasi (razmak između svakog slova):
10 110 111 10 110 111 0 111 10 0

b) Binarni Huffmanov kod sa simbolima 0 i 1:

0.4 B	0.4 B	D/0 0.6	**D/00 1**
0.35 D	0.35 D	A/10	**A/010**
0.15 A	A/0 0.25	C/11	**C/011**
0.1 C	C/1	0.4 B	**B/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:

$$\overline{I(X)} = \frac{H(X/X^{\infty})}{n_{sr} \cdot \tau} = \frac{0.9738}{\tau}$$

odnosno iskorištenost kanala veze je 97.38%.

Kodirana poruka DACDACBCDB glasi (razmak između svakog slova):
00 010 011 00 010 011 1 011 00 1

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

BB 0,16		**0,16/00**				
BD 0,14	0,44/0	0,28/01	**0,14/010**			
DB 0,14			**0,14/011**			
DD 0,1225			**0,1225/100**			
AB 0,06		0,295/10		**0,06/1010**		
BA 0,06			0,1725/101	0,1125/1011	**0,06/10110**	
AD 0,0525					**0,0525/10111**	
DA 0,0525				**0,0525/1100**		
BC 0,04			0,1325/110	0,08/1101	**0,04/11010**	
CB 0,04	0,56/1				**0,04/11011**	
CD 0,035				0,07/1110	**0,035/11100**	
DC 0,035		0,265/11			**0,035/11101**	
AA 0,0225			0,1325/111		0,0375/11110	**0,0225/111100**
AC 0,015				0,0625/1111		**0,015/111101**
CA 0,015					0,025/11111	**0,015/111110**
CC 0,01						**0,01/111111**

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

$$n_{sr} = 0.16 \cdot 2 + 0.14 \cdot 3 + 0.14 \cdot 3 + 0.1225 \cdot 3 + 0.06 \cdot 4 + 0.06 \cdot 5 + 0.0525 \cdot 5 + 0.0525 \cdot 4 + 0.04 \cdot 5 + 0.04 \cdot 5$$

$$+0.035 \cdot 5 + 0.035 \cdot 5 + 0.0225 \cdot 6 + 0.015 \cdot 6 + 0.015 \cdot 6 + 0.01 \cdot 6 = 3.665$$

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

$$\overline{I(X)} = \frac{2 \cdot H(X/X^{\infty})}{n_{sr} \cdot \tau} = \frac{0.9831}{\tau}$$

odnosno iskorištenost kanala veze je 98.31%.
Kodirana poruka DACDACBCDB glasi (razmak između svaka 2 slova):
1100 11100 111101 11010 011

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

BB 0,16	BB 0,16	BB 0,16	BB 0,16	BB 0,16	BB 0,16
BD 0,14	BD 0,14	BD 0,14	BD 0,14	BD 0,14	BD 0,14
DB 0,14	DB 0,14	DB 0,14	DB 0,14	DB 0,14	DB 0,14
DD 0,1225	DD 0,1225	DD 0,1225	DD 0,1225	DD 0,1225	DD 0,1225
AB 0,06	AB 0,06	AB 0,06	AB 0,06	AA/00 0,0725	BC/0 0,08
BA 0,06	BA 0,06	BA 0,06	BA 0,06	AC/01	CB/1
AD 0,0525	AD 0,0525	AD 0,0525	DC/0 0,06	CD/1	AA/00 0,0725
DA 0,0525	DA 0,0525	DA 0,0525	CA/10	AB 0,06	AC/01
BC 0,04	BC 0,04	BC 0,04	CC/11	BA 0,06	CD/1
CB 0,04	CB 0,04	CB 0,04	AD 0,0525	DC/0 0,06	AB 0,06
CD 0,035	CD 0,035	AA/0 0,0375	DA 0,0525	CA/10	BA 0,06
DC 0,035	DC 0,035	AC/1	BC 0,04	CC/11	DC/0 0,06
AA 0,0225	CA/0 0,025	CD 0,035	CB 0,04	AD 0,0525	CA/10
AC 0,015	CC/1	DC 0,035	AA/0 0,0375	DA 0,0525	CC/11
CA 0,015	AA 0,0225	CA/0 0,025	AC/1	BC 0,04	AD 0,0525
CC 0,01	AC 0,015	CC/1	CD 0,035	CB 0,04	DA 0,0525

BB 0,16	BB 0,16	BB 0,16	AD/00 0,185	DD/0 0,2425
BD 0,14	BD 0,14	BD 0,14	DA/01	BA/10
DB 0,14	DB 0,14	DB 0,14	BC/10	DC/110
DD 0,1225	DD 0,1225	AA/000 0,1325	CB/11	CA/1110
AD/0 0,105	BA/0 0,12	AC/001	BB 0,16	CC/1111
DA/1	DC/10	CD/01	BD 0,14	AD/00 0,185
BC/0 0,08	CA/110	AB/1	DB 0,14	DA/01
CB/1	CC/111	DD 0,1225	AA/000 0,1325	BC/10
AA/00 0,0725	AD/0 0,105	BA/0 0,12	AC/001	CB/11
AC/01	DA/1	DC/10	CD/01	BB 0,16
CD/1	BC/0 0,08	CA/110	AB/1	BD 0,14
AB 0,06	CB/1	CC/111	DD 0,1225	DB 0,14
BA 0,06	AA/00 0,0725	AD/0 0,105	BA/0 0,12	AA/000 0,1325
DC/0 0,06	AC/01	DA/1	DC/10	AC/001
CA/10	CD/1	BC/0 0,08	CA/110	CD/01
CC/11	AB 0,06	CB/1	CC/111	AB/1

.

DB/0 0,2725	BB/0 0,3	DD/00 0,4275	BB/00 0,5725	**BB/000 1**
AA/1000	BD/1	BA/010	BD/01	**BD/001**
AC/1001	DB/0 0,2725	DC/0110	DB/10	**DB/010**
CD/101	AA/1000	CA/01110	AA/11000	**AA/011000**
AB/11	AC/1001	CC/01111	AC/11001	**AC/011001**
DD/0 0,2425	CD/101	AD/100	CD/1101	**CD/01101**
BA/10	AB/11	DA/101	AB/111	**AB/0111**
DC/110	DD/0 0,2425	BC/110	DD/00 0,4275	**DD/100**
CA/1110	BA/10	CB/111	BA/010	**BA/1010**
CC/1111	DC/110	BB/0 0,3	DC/0110	**DC/10110**
AD/00 0,185	CA/1110	BD/1	CA/01110	**CA/101110**
DA/01	CC/1111	DB/0 0,2725	CC/01111	**CC/101111**
BC/10	AD/00 0,185	AA/1000	AD/100	**AD/1100**
CB/11	DA/01	AC/1001	DA/101	**DA/1101**
BB 0,16	BC/10	CD/101	BC/110	**BC/1110**
BD 0,14	CB/11	AB/11	CB/111	**CB/1111**

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

$$n_{sr} = 3 \cdot 0.16 + 3 \cdot 0.14 + 3 \cdot 0.14 + 6 \cdot 0.0225 + 6 \cdot 0.015 + 5 \cdot 0.035 + 4 \cdot 0.06 + 3 \cdot 0.1225$$

$$+4 \cdot 0.06 + 5 \cdot 0.035 + 6 \cdot 0.015 + 6 \cdot 0.01 + 4 \cdot 0.0525 + 4 \cdot 0.0525 + 4 \cdot 0.04 + 4 \cdot 0.04 = 3.6325$$

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

$$\overline{I(X)} = \frac{2 \cdot H(X/X^{\infty})}{n_{sr} \cdot \tau} = \frac{0.99193}{\tau}$$

odnosno iskorištenost kanala veze je 99.193%.
Kodirana poruka DACDACBCDB glasi (razmak između svaka 2 slova):
1101 01101 011001 1110 010

7. Rješenje zadatka

Zadatak 7 [0.8 poena]

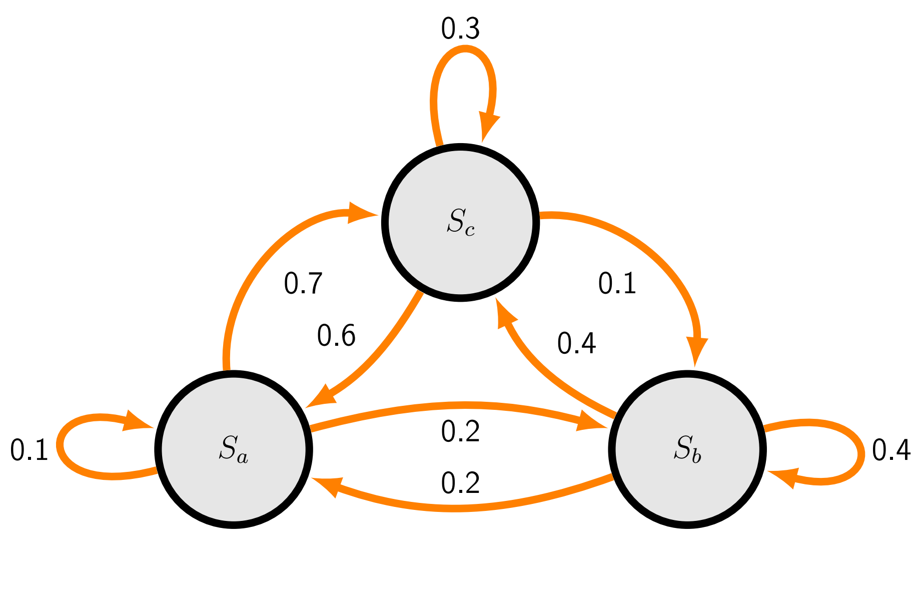
Markovljev izvor informacija prvog reda emitira tri različite poruke a, b i c. Ovisno od toga koja je poruka posljednja emitirana, izvor se nalazi u jednom od 3 moguća stanja S_a, S_b i S_c koja redom odgovaraju emitiranim porukama a, b odnosno c. Vjerovatnoće da će izvor emitirati neku od ove 3 poruke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

p(x_j /S_i)	a	b	c
S_a	0.1	0.2	0.7
S_b	0.2	0.4	0.4
S_c	0.6	0.1	0.3

Za ovaj izvor informacija formirajte binarni Shannon-Fano kod sa simbolima 0 i 1
a. posmatrajući izvor kao izvor bez memorije;
b. posmatrajući izvor kao izvor bez memorije, ali kodirajući parove poruka umjesto individualnih poruka;
c. koristeći posebno kodiranje za svako stanje;
d. 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 cbbcbaba. 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:

$$P(S_a) = 0.3486239 \quad P(S_b) = 0.1926606 \quad P(S_c) = 0.4587156$$

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

$$H(S_a) = 1.15678 \quad H(S_b) = 1.52193 \quad H(S_c) = 1.29546$$

Sada možemo izračunati entropiju izvora:

$$H(X/X^{\infty}) = P(S_a)H(S_a) + P(S_b)H(S_b) + P(S_c)H(S_c) = 1.290744813176$$

- Pređimo sada na kodiranja (S_a -> a, S_b -> b, S_c -> c):

a) posmatrajući izvor kao izvor bez memorije:

c 0.4587156	**0.4587156/0**	
a 0.3486239	0.5412845/1	**0.3486239/10**
b 0.1926606		**0.1926606/11**

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

$$n_{sr} = 0.4587156 + 0.3486239 \cdot 2 + 0.1926606 \cdot 2 = 1.5412846$$

samim tim protok je:

$$\overline{I(X)} = \frac{H(X/X^{\infty})}{n_{sr} \cdot \tau} = \frac{0.837447}{\tau}$$

odnosno iskorištenost je: 83.74%.
Kodirana poruka cbbcbaba glasi (razmak između svakog slova):
0 11 11 0 11 10 11 10

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

cc 0.21042000168	0.458/0	**0.21042000168/00**		
ac 0.15991922146		0.319838442925/01	**0.15991922146/010**	
ca 0.15991922146			**0.15991922146/011**	
aa 0.12153862365	0.541/1	0.20991504637/10	**0.12153862365/100**	
bc 0.08837642272			**0.08837642272/101**	
cb 0.08837642272		0.2598267090/11	0.15554251246/110	**0.08837642272/1100**
ab 0.06716608974				**0.06716608974/1101**
ba 0.06716608974			0.10428419656/111	**0.06716608974/1110**
bb 0.03711810679				**0.03711810679/1111**

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

$$n_{sr} = 0.21042000168336 \cdot 2 + 0.15991922146284 \cdot 3 + 0.15991922146284 \cdot 3 + 0.12153862365121 \cdot 3+$$

$$+0.08837642272536 \cdot 3 + 0.08837642272536 \cdot 4+$$

$$0.06716608974834 \cdot 4 + 0.06716608974834 \cdot 4 + 0.03711810679236 \cdot 4 = 3.04940730733107$$

samim tim protok je:

$$\overline{I(X)} = \frac{2 \cdot H(X/X^{\infty})}{n_{sr} \cdot \tau} = \frac{0.846554548}{\tau}$$

odnosno iskorištenost je: 84.65%.
Kodirana poruka cbbcbaba glasi (razmak između svaka 2 slova):
1100 101 1110 1110

c) koristeći posebno kodiranje za svako stanje:

S_a

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

$$n_{sr_a} = 0.7 + 0.4 + 0.2 = 1.3$$

S_b

c 0.4	**0.4/0**	
b 0.4	0.6/1	**0.4/10**
a 0.2		**0.2/11**

$$n_{sr_b} = 0.4 + 0.8 + 0.4 = 1.6$$

S_c

a 0.6	**0.6/0**	
c 0.3	0.4/1	**0.3/10**
b 0.1		**0.1/11**

$$n_{sr_c} = 0.6 + 0.6 + 0.2 = 1.4$$

$$n_{sr} = P(S_a) \cdot n_{sr_a} + P(S_b) \cdot n_{sr_b} + P(S_c) \cdot n_{sr_c} = 0.3486239 \cdot 1.3 + 0.1926606 \cdot 1.6 + 0.4587156 \cdot 1.4$$

$$n_{sr} = 1.40366987$$

Pošto izvor započinje rad u stanju S_a kodirana poruka cbbcbaba glasi (razmak po slovu):
0 11 10 0 11 11 10 11

Protok informacija je:

$$\overline{I(X)} = \frac{H(X/X^{\infty})}{n_{sr} \cdot \tau} = \frac{0.919550}{\tau}$$

Iskorištenost kanala veze je približno 91.955%.

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

S_a

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

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

S_b

cc 0.16		**0.16/00**		
cb 0.16	0.48/0	0.32/01	**0.16/010**	
bc 0.16			**0.16/011**	
bb 0.16		0.24/10	**0.16/100**	
ca 0.08			**0.08/101**	
ac 0.08	0.52/1		0.16/110	**0.08/1100**
ab 0.08		0.28/11		**0.08/1101**
ba 0.08			0.12/111	**0.08/1110**
aa 0.04				**0.04/1111**

$$n_{sr_b} = 0.16 \cdot 2 + 0.16 \cdot 3 + 0.16 \cdot 3 + 0.16 \cdot 3 + 0.08 \cdot 3 + 0.08 \cdot 4 + 0.08 \cdot 4 + 0.08 \cdot 4 + 0.04 \cdot 4 = 3.12$$

S_c

aa 0.36	0.54/0	**0.36/00**			
ac 0.18		**0.18/01**			
ca 0.18		0.27/10	**0.18/100**		
cc 0.09			**0.09/101**		
ba 0.06	0.46/1		0.12/110	**0.06/1100**	
ab 0.06				**0.06/1101**	
cb 0.03		0.19/11		**0.03/1110**	
bc 0.03			0.07/111	0.04/1111	**0.03/11110**
bb 0.01					**0.01/11111**

$$n_{sr_c} = 0.36 \cdot 2 + 0.18 \cdot 2 + 0.18 \cdot 3 + 0.09 \cdot 3 + 0.06 \cdot 4 + 0.06 \cdot 4 + 0.03 \cdot 4 + 0.03 \cdot 5 + 0.01 \cdot 5 = 2.69$$

Izračunajmo n_{sr}:

$$n_{sr} = P(S_a) \cdot n_{sr_a} + P(S_b) \cdot n_{sr_b} + P(S_c) \cdot n_{sr_c} = 0.3486239 \cdot 2.33 + 0.1926606 \cdot 3.12 + 0.4587156 \cdot 2.69$$

$$n_{sr} = 2.647339723$$

Protok informacija je:

$$\overline{I(X)} = \frac{2 \cdot H(X/X^{\infty})}{n_{sr} \cdot \tau} = \frac{0.97512}{\tau}$$

Iskorištenost kanala veze je približno 97.512%.

Pošto izvor započinje rad u stanju S_a kodirana poruka cbbcbaba glasi (razmak po svakom markovljevom lancu):
100 011 1100 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 0.05 i 0.2 respektivno. Odredite količinu prenesene informacije kroz ovaj kanal ukoliko vjerovatnoća pojave nule na ulazu u kanal iznosi 0.5, 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:

$$p(z_1/y_1) = 0.95 \quad p(z_2/y_1) = 0.05 \quad p(z_1/y_2) = 0.2 \quad p(z_2/y_2) = 0.8$$

odnosno:

$$p(y_1) = 0.5 \quad p(y_2) = 0.5$$

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

$$p(z_1) = p(y_1) \cdot p(z_1/y_1) + p(y2) \cdot p(z_1/y_2) = 0.5 \cdot 0.95 + 0.5 \cdot 0.2 = 0.575$$

$$p(z_2) = 1 - p(z_1) = 1 - 0.575 = 0.425$$

Za entropije ulaznih i izlaznih simbola imamo:

$$H(Y) = -(p(y_1) \cdot log_2\, p(y_1) + p(y_2) \cdot log_2\, p(y_2)) = 1$$

$$H(Z) = -(p(z_1) \cdot log_2\, p(z_1) + p(z_2) \cdot log_2\, p(z_2)) = 0.983708$$

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):

$$H(Z/y_1) = -(p(z_1/y_1) \cdot log2(p(z_1/y_1)) + p(z_2/y_1) \cdot log2(p(z_2/y_1)) = 0.286397$$

$$H(Z/y_2) = -(p(z_1/y_2) \cdot log2(p(z_1/y_2)) + p(z_2/y_2) \cdot log2(p(z_2/y_2)) = 0.721928$$

Odavde za uvjetnu entropiju H(Z/Y) dobijamo:

$$H(Z/Y) = p(y_1) \cdot H(Z/y_1) + p(y_2) \cdot H(Z/y_2) = 0.5 \cdot 0.286397 + 0.5 \cdot 0.721928 = 0.5041625$$

Prenesena količina informacije kroz kanal je data sa:

$$I(Y, Z) = H(Z) - H(Z/Y) = 0.983708 - 0.5041625 = 0.4795455$$

Kapacitet kanala veze je dat relacijom:

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

gdje su:

$$\alpha = 0.05 \quad \beta = 0.2 \quad H_0 = 0.286397 \quad H_1 = 0.721928 \quad \gamma = 1 - \alpha - \beta = 0.75$$

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

$$C_c = \frac{0.481307}{\tau}$$

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) = 0.3. 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 0.1 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:

a. Prosto kodiranje a -> 0 i b -> 1 uz dekodiranje zasnovano na klasifikaciji $S_a = \{0\}$ i $S_b = \{1\}$;
b. Zaštitno kodiranje a -> 000 i b -> 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:

$$p(z_1/y_1) = 0.75 \quad p(z_2/y_1) = 0.25 \quad p(z_1/y_2) = 0.1 \quad p(z_2/y_2) = 0.9$$

odnosno:

$$p(a) = 0.3 \quad p(b) = 0.7$$

Neka su poruke koje prima krajnji korisnik w_1 i w_2.

a) Kako se kodiranje vrši prostim ravnomjernim kodom a -> 0 i b -> 1 (tj. a -> y_1 i b -> y_2) uz dekodiranje zasnovano na klasifkaciji $S_a = \{0\}$ 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:

$$p(w_1/a) = p(z_1/y_1) = 0.75$$
$$p(w_2/a) = 1 - p(w_1/a) = 0.25$$
$$p(w_1/b) = p(z_1/y_2) = 0.1$$
$$p(w_2/b) = 1 - p(w_1/b) = 0.9$$

Dalje imamo:

$$p(w_1) = p(a) \cdot p(w_1/a) + p(b) \cdot p(w_1/b) = 0.3 \cdot 0.75 + 0.7 \cdot 0.1 = 0.295$$

$$p(w_2) = 1 - p(w_1) = 0.705$$

Sada možemo izračunati naredno:

$$H(W) = -(p(w_1) \cdot log_2\ p(w_1) + p(w_2) \cdot log_2\ p(w_2)) = 0.875093$$
$$H(W/a) = -(p(w_1/a) \cdot log_2\ p(w_1/a) + p(w_2/a) \cdot log_2\ p(w_2/a)) = 0.811278$$
$$H(W/b) = -(p(w_1/b) \cdot log_2\ p(w_1/b) + p(w_2/b) \cdot log_2\ p(w_2/b)) = 0.468996$$
$$H(W/X) = p(a) \cdot H(W/a) + p(b) \cdot H(W/b) = 0.5716806$$
$$I(X,W) = H(W) - H(W/X) = 0.875093 - 0.5716806 = 0.3034124$$

Odredimo brzinu prenosa informacija kroz komunikacioni kanal:

$$\overline{I(X,W)} = \frac{I(X,W)}{\tau} = \frac{0.3034124}{\tau}$$

Procenat iskorištenosti kanala veze je 88.1326% (iskorištenost u prenosu / C_c - C_c dobijemo po ogromnoj formuli iz 8 zadatka). Izračunajmo još vjerovatnoću greške u prenosu:

$$p_e = 1 - p(a) \cdot p(w_1/a) - p(b) \cdot p(w_2/b) = 1 - 0.3 \cdot 0.75 - 0.7 \cdot 0.9 = 0.145$$

b) Kodiranje se vrši zaštitnim kodom a -> 000 i b -> 111 (tj. a -> $y_1y_1y_1$ i b -> $y_2y_2y_2$) 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:

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

Na isti način dobijamo da je

$$p(w_1/b) = p(z_1/y_2)^3 + 3 \cdot (z_1/y_2)^2 \cdot p(z_2/y_2) = 0.1^3 + 3 \cdot 0.1^2 \cdot 0.9 = 0.028$$
$$p(w_2/b) = 1 - 0.028 = 0.972$$

Dalje imamo:

$$p(w_1) = p(a) \cdot p(w_1/a) + p(b) \cdot p(w_1/b) = 0.3 \cdot 0.84375 + 0.7 \cdot 0.15625 = 0.362465$$

$$p(w_2) = 1 - p(w_1) = 0.637535$$

Sada možemo izračunati naredno:

$$H(W) = -(p(w_1) \cdot log_2\ p(w_1) + p(w_2) \cdot log_2\ p(w_2)) = 0.944710$$
$$H(W/a) = -(p(w_1/a) \cdot log_2\ p(w_1/a) + p(w_2/a) \cdot log_2\ p(w_2/a)) = 0.625262$$
$$H(W/b) = -(p(w_1/b) \cdot log_2\ p(w_1/b) + p(w_2/b) \cdot log_2\ p(w_2/b)) = 0.184261$$
$$H(W/X) = p(a) \cdot H(W/a) + p(b) \cdot H(W/b) = 0.3165613$$
$$I(X,W) = H(W) - H(W/X) = 0.944710 - 0.3165613 = 0.6281487$$

Odredimo brzinu prenosa informacija kroz komunikacioni kanal:

$$\overline{I(X,W)} = \frac{I(X,W)}{3 \cdot \tau} = \frac{0.6281487}{3 \cdot \tau} = \frac{0.20938}{\tau}$$

Procenat iskorištenosti kanala veze je 35.86% (iskorištenost u prenosu / C_c - C_c dobijemo po ogromnoj formuli iz 8 zadatka). Izračunajmo još vjerovatnoću greške u prenosu:

$$p_e = 1 - p(a) \cdot p(w_1/a) - p(b) \cdot p(w_2/b) = 1 - 0.3 \cdot 0.84375 - 0.7 \cdot 0.972 = 0.066475$$