Sveučilište u Zagrebu
Fakultet elektrotehnike i računarstva

2. ispitni rok (jesenski) iz predmeta **TEORIJA INFORMACIJE**, 27. kolovoza 2020.

Pravilo bodovanja zadataka

Svaki točno odgovoreni zadatak (osim 7. i 8.) donosi 10 bodova, netočno odgovoreni 4 negativna boda, a neodgovoreni 0 bodova. Točno odgovoreni 7. i 8. zadatak donose po 20 bodova, netočno odgovoreni 8 negativnih boda, a neodgovoreni 0 bodova.

Zadatak 1. (10 bodova) Bezmemorijski informacijski izvor generira simbole iz skupa X = {a, b, c, d, e, f, g} s vjerojatnostima pojavljivanja P(a) = 0,22, P(b) = 0,35, P(c) = 0,15, P(d) = P(e) = 0,09, P(f) = P(g) = 0,05. Dani skup simbola kodiran je binarnom tehnikom Shannon-Fano. Odredite efikasnost koda, a rezultat zaokružite na dvije decimale.

a) 0,97

b) 0,99

c) 0,96

d) 0,98

e) ništa od navedenog.

Postupak rješavanja:

Provedbom Shannon-Fanoovg kodiranja dobivamo sljedeće kodne riječi:

b	00
a	01
c	100
d	101
e	110
f	1110
g	1111

Iz dobivenih kodnih riječi očitavamo njihove duljine l_i te određujemo srednju duljinu kodne riječi, L:

$$L = \sum_{i=1}^{n} P_i \cdot l_i = 2,53 \frac{\text{bit}}{\text{simbol}},$$

pri čemu je P_i skraćena oznaka za vjerojatnosti simbola iz skupa X. Efikasnost koda određena je izrazom

$$\varepsilon = \frac{H(X)}{L} .$$

S obzirom da vrijedi

$$H\left(X\right) = -\sum_{i=1}^{n} P_i \log_2 P_i = 2,4787 \frac{\text{bit}}{\text{simbol}},$$

dobivamo da je ε = 0,9797 ≈ 0,98.

Zadatak 2. (10 bodova) Zadana su dva paralelna kanala u kojima djeluje aditivni bijeli Gaussov šum Z_1, odnosno Z_2, oba s očekivanjem nula. Također vrijedi: $E[Z_1^2] = 0{,}5$, odnosno $E[Z_2^2] = 0{,}7$. Na ulazu prvog kanala djeluje slučajni signal X_1, a na ulazu drugog kanala slučajni signal X_2. Neka je $E[X_1] = E[X_2] = 0$ te $E[X_1^2] + E[X_2^2] = 0.4$. Odredite maksimalnu dinamiku u zadanom sustavu kanala (bit/simbol). Napomena: pri određivanju ekstrema dinamike na promatranom intervalu potrebno je načiniti analizu vrijednosti koje dinamika postiže u točkama ekstrema i temeljem toga zaključiti koja je vrijednost maksimalna. Nije potrebno provoditi drugu derivaciju zbog prevelike složenosti izraza.

a) 0,87 bit/simbol

b) 0,435 bit/simbol

c) 1 bit/simbol

d) 0,326 bit/simbol

e) ništa od navedenog.

Postupak rješavanja:

Dinamika u sustavu paralelnih kanala jednaka je zbroju dinamika pojedinih kanala, tj.:

$$D = D_1 + D_2 = \frac{1}{2}\log_2\left(1+\frac{\sigma_{X_1}^2}{\sigma_{Z_1}^2}\right) + \frac{1}{2}\log_2\left(1+\frac{\sigma_{X_2}^2}{\sigma_{Z_2}^2}\right) = \frac{1}{2}\log_2\left(1+\frac{\sigma_{X_1}^2}{\sigma_{Z_1}^2}\right) + \frac{1}{2}\log_2\left(1+\frac{0,4-\sigma_{X_1}^2}{\sigma_{Z_2}^2}\right)$$

Maksimum dinamike nalazimo deriviranjem prethodnog izraza po σ_1^2 i izjednačavanjem istog s nulom. Dakle, iz

$$\frac{dD}{d\sigma_{X_1}} = 0$$

deriviranjem dobivamo (skraćeni prikaz konačnog uvjeta):

$$2\sigma_{X_1}\left(0,3-\sigma_{X_1}^2\right) = 0$$

Dakle, ekstremi su:

$$\left(\sigma_{X_1}\right)_1 = 0,\ \left(\sigma_{X_1}\right)_2 = \sqrt{0,3}\,.$$

Za prvu vrijednost dobivamo D_1 = 0,326 bit/simbol, a za drugu D_2 = 0,435 bit/simbol. Očito je da maksimum dinamike nastupa za drugu vrijednost, pa je konačan rezultat D_{max} = 0,435 bit/simbol.

Zadatak 3. (10 bodova) U nekom komunikacijskom sustavu abeceda izvora sadrži 5 simbola, x_1, x_2, x_3, x_4 i x_5, a abeceda odredišta 4 simbola, y_1, y_2, y_3 i y_4. Matrica združenih vjerojatnosti simbola izvora i odredišta zadana je izrazom:

$$\left[P(X,Y)\right] = \begin{bmatrix} 0,25 & 0 & 0 & 0 \\ 0,1 & 0,3 & 0 & 0 \\ 0 & 0,05 & 0,1 & 0 \\ 0 & 0 & 0,05 & 0,1 \\ 0 & 0 & 0,05 & 0 \end{bmatrix}$$

Odredite ekvivokaciju u kanalu koji povezuje gore opisani izvor i odredište.

a) 0,809 bit/simbol

b) 2,667 bit/simbol

c) 0,667 bit/simbol

d) 1,667 bit/simbol

e) ništa od navedenog.

Postupak rješavanja:

Temeljem zadane matrice združenih vjerojatnosti simbola zbrajanjem vjerojatnosti po recima, odnosno po stupcima možemo dobiti vjerojatnosti $p(x_i)$, odnosno $p(y_j)$:

$p(x_1) = 0,25$

$p(x_2) = 0,1 + 0,3 = 0,4$

$p(x_3) = 0,05 + 0,1 = 0,15$

$p(x_4) = 0,05 + 0,1 = 0,15$

$p(x_5) = 0,05$

Provjera: $\sum_{i=1}^{5} p\left(x_i\right) = 0,25 + 0,4 + 0,15 + 0,15 + 0,05 = 1$

$p(y_1) = 0,25 + 0,1 = 0,35$

$p(y_2) = 0,3 + 0,05 = 0,35$

$p(y_3) = 0,1 + 0,05 + 0,05 = 0,2$

$p(y_4) = 0,1$

Provjera: $\sum_{j=1}^{4} p\left(y_j\right) = 0,35 + 0,35 + 0,2 + 0,1 = 1$

Ekvivokaciju $H(X|Y)$ računamo prema sljedećem izrazu:

$$H\left(X|Y\right) = -\sum_{i=1}^{5}\sum_{j=1}^{4} p\left(x_i, y_j\right) \cdot \log_2\left[\frac{p\left(x_i, y_j\right)}{p\left(y_j\right)}\right] \text{[bit/simbol]}$$

Uvrštavanjem gore zadanih te izračunatih vrijednosti dobivamo $H(X|Y) = 0,809$ bit/simbol.

Zadatak 4. (10 bodova) Zadan je binarni blok kôd K. Na ulazu kodera kanala danog koda pojavljuju se tri poruke: $\mathbf{d}_1 = [101]$, $\mathbf{d}_2 = [011]$ i $\mathbf{d}_3 = [111]$. Na izlazu kodera kanala, za dane tri poruke $\mathbf{d}_1$, $\mathbf{d}_2$ i $\mathbf{d}_3$ pojavljuju se sljedeće tri kodne riječi: $\mathbf{c}_1 = [100101]$, $\mathbf{c}_2 = [001011]$, odnosno $\mathbf{c}_3 = [010110]$. Odredite 5. i 6. bit u kodnoj riječi koja odgovara poruci $\mathbf{d}_4 = [110]$.

a) 00

b) 10

c) 01

d) 11

Postupak rješavanja:

Temeljem zadanih poruka i kodnih riječi očito je da se radi o kodu $[n, k] = [6, 3]$. Također, temeljem zadanih kodnih riječi $\mathbf{c}_1$, $\mathbf{c}_2$ i $\mathbf{c}_3$ vidljivo je da generirajuća matrica danog koda K nije u standardnom obliku. Dakle, neka je

$$\mathbf{G} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \\ c_1 & c_2 & c_3 & c_4 & c_5 & c_6 \end{bmatrix}, a_i, b_i, c_i \in \{0,1\} \; za \; i = 1, \ldots, 6$$

Nadalje, za zadane poruke i kodne riječi vrijede sljedeće jednakosti:

$$\mathbf{d}_1 \cdot \mathbf{G} = \mathbf{c}_1 \rightarrow [101] \cdot \mathbf{G} = [100101] \qquad (1)$$

$$\mathbf{d}_2 \cdot \mathbf{G} = \mathbf{c}_2 \rightarrow [011] \cdot \mathbf{G} = [001011] \qquad (2)$$

$$\mathbf{d}_3 \cdot \mathbf{G} = \mathbf{c}_3 \rightarrow [111] \cdot \mathbf{G} = [010110] \qquad (3)$$

Temeljem navedenih jednakosti određujemo bitove matrice **G**. Na primjer,

Temeljem jednakosti (1) za prvi bit kodne riječi $\mathbf{c}_1$ vrijedi sljedeće: $a_1 \oplus c_1 = 1$

Temeljem jednakosti (2) za prvi bit kodne riječi $\mathbf{c}_2$ vrijedi sljedeće: $b_1 \oplus c_1 = 0$

Temeljem jednakosti (3) za prvi bit kodne riječi $\mathbf{c}_3$ vrijedi sljedeće: $a_1 \oplus b_1 \oplus c_1 = 0$

Iz toga slijedi: $a_1 = 0$, $b_1 = 1$ i $c_1 = 1$. Po istoj analogiji određujemo sve ostale bitove matrice **G** i dobivamo:

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Dakle, $\mathbf{c}_4 = \mathbf{d}_4 \cdot \mathbf{G} = [101110]$, odnosno bitovi koji odgovaraju 5. i 6. poziciji u kodnoj riječi $\mathbf{c}_4$ su 1 i 0.

Zadatak 5. (10 bodova) Koder kanala u nekom komunikacijskom sustavu koristi Hammingov kôd zadan matricom provjere pariteta Ham(r). Odredite koliko najmanje mora iznositi r pa da kodna brzina ovako zadanog linearnog binarnog blok koda bude veća od 0,904.

a) 5 bita

b) 4 bita

c) 6 bita

d) 7 bita

e) ništa od navedenog.

Postupak rješavanja:

Hammingov kôd definiran matricom Ham(r) koristi r zaštitnih bita koji štite m bita poruke. Kodna brzina $R(K)$ definirana je kao omjer broja bita poruke, m, prema ukupnom broju bita u kodnoj riječi (zbroj broja bita poruke i broja zaštitnih bita), $m + r$. Nadalje, vrijedi relacija:

$m + r \leq 2^r - 1$. Dakle, kodna brzina uz zadani broj zaštitnih bita bit će maksimalna kad vrijedi

$m + r = 2^r - 1$. Sukladno tome, kodna brzina ovog koda Ham(r) bit će jednaka:

$$R(K)=\frac{2^r-1-r}{2^r-1}=1-\frac{r}{2^r-1}$$

Uzevši u obzir uvjet $R(K) > 0{,}904$ i ponuđena rješenja dobivamo da je uvjet ispunjen uz $r = 6$ bita.

Zadatak 6. (10 bodova) Zadan je diskretni binarni kanal. Na izvoru informacije pojavljuju se simboli x_1 i x_2, a na odredištu simboli y_1 i y_2. Matrica šuma u kanalu koji povezuje izvor i odredište, $[P(Y|X)]$, zadana je kao

$$P[Y \mid X]=\begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix}.$$

Odredite za koliko se kapacitet ovog zadanog kanala razlikuje od maksimalnog mogućeg kapaciteta binarnog simetričnog kanala (traži se apsolutna vrijednost razlike).

a) 0 bit/simbol

b) 1 bit/simbol

c) 0,92 bit/simbol

d) 0,08 bit/simbol

e) ništa od navedenog.

Postupak rješavanja:

U zadanom kanalu vjerojatnost pogrešnog prijenosa simbola je $p_g = 1/3$. Kapacitet binarnog simetričnog kanala dan je izrazom:

$$C=1+p_g\log_2(p_g)+(1-p_g)\log_2(1-p_g)[\text{bit/simbol}]$$

Ako u izraz uvrstimo $p_g = 1/3$, dobit ćemo $C = 0{,}08$ bit/simbol. Maksimalan mogući kapacitet binarnog simetričnog kanala, C_{max}, iznosi 1 bit/simbol. Sukladno tome, apsolutna vrijednost razlike između C_{max} i C iznosi 0,92 bit/simbol (rješenje označeno slovom d).

Zadatak 7. (20 bodova) Razmatrajte izvor koji generira četiri simbola iz skupa $X = \{x_1, x_2, x_3, x_4\}$ s odgovarajućim vjerojatnostima pojavljivanja za koje vrijedi:

$$1 > p(x_1) = p_1 > p(x_2) = p_2 > p(x_3) = p_3 > p(x_4) = p_4 > 0 \text{ i } \sum_{i=1}^{4} p_i = 1.$$

Svi su simboli potpuno neovisni jedni o drugima. Nadalje, izvor je spojen s koderom informacije koji navedene simbole kodira binarnim simbolima sukladno algoritmu Shannon-Fano, a rezultat toga je prefiksni kôd. Kodne riječi na izlazu kodera informacije, $C(x_i)$, ovise o razdiobi vjerojatnosti simbola $x_i \in X$. Neka su zadane vjerojatnosti $p_3 = 0{,}19$ i $p_4 = 0{,}15$. Odredite granice unutar kojih se smije nalaziti p_1 pa da kodna riječ $C(x_1)$ može imati duljinu jedan bit.

a) [0,5, 0,66)

b) [0,5, 1)

c) [0,34, 0,47]

d) [0,34, 0,47)

e) ništa od navedenog.

Postupak rješavanja:

a) Način kodiranja algoritmom Shannon-Fano ovisi o razdiobi vjerojatnosti $p(x_i)$. Pri tome je važno kako se simboli x_i, ovisno o $p(x_i)$, grupiraju. Bit je algoritma da prilikom podjele simbola u dvije grupe razlika zbroja vjerojatnosti simbola u jednoj i drugoj grupi bude minimalna. U slučaju zadanih simbola x_i i adekvatne razdiobe vjerojatnosti $p(x_i)$, konačan rezultat kodiranja algoritmom Shannon-Fano može biti:

1) $C(x_1) = 00$, $C(x_2) = 01$, $C(x_3) = 10$, $C(x_4) = 11$, ili

2) $C(x_1) = 0$, $C(x_2) = 10$, $C(x_3) = 110$, $C(x_4) = 111$.

Dakle, samo u drugom ishodu kodiranja moguće je ostvariti da $C(x_1)$ ima duljinu jednog bita. Da bi se simboli x_i dijelili u grupe na način koji odgovara binarnom kodu kreiranom u ishodu 2, mora vrijediti:

$| p_1 - (p_2 + p_3 + p_4)| \leq | (p_1 + p_2) - (p_3 + p_4)|$, tj. s obzirom da je $p_3 + p_4 = 0{,}19 + 0{,}15 = 0{,}34$

$| p_1 - p_2 - 0{,}34 | \leq | p_1 + p_2 - 0{,}34|$

Desna strana nejednakosti uvijek vrijedi zbog uvjeta $1 > p(x_1) > p(x_2) > p(x_3) > p(x_4) > 0$. Lijeva strana nejednakosti će polučiti sljedeći rezultat:

- za $p_1 \geq p_2 + p_3 + p_4$ vrijedi: $p_1 - p_2 - p_3 - p_4 \leq p_1 + p_2 - p_3 - p_4$, što daje: $2p_2 \geq 0$, a to uvijek vrijedi;

Međutim, iz uvjeta $p_1 \geq p_2 + p_3 + p_4$, tj. $p_1 \geq p_2 + 0{,}34$, te uz $p_2 = 1 - p_1 - (p_3 + p_4) = 0{,}66 - p_1$ mora vrijediti: $2p_1 \geq 1$, tj. $\boldsymbol{p_1 \geq 0{,}5}$; istovremeno, zbog uvjeta $p_2 > p_3$, tj. $p_2 > 0{,}19$, te zbog jednakosti $p_1 + p_2 = 1 - (p_3 + p_4)$, slijedi $p_1 < 0{,}66 - 0{,}19$, tj. $\boldsymbol{p_1 < 0{,}47}$. S obzirom da je ova dva uvjeta za p_1 nemoguće istovremeno zadovoljiti, $p_1 \geq p_2 + p_3 + p_4$ nije opcija koja pogoduje rješenju.

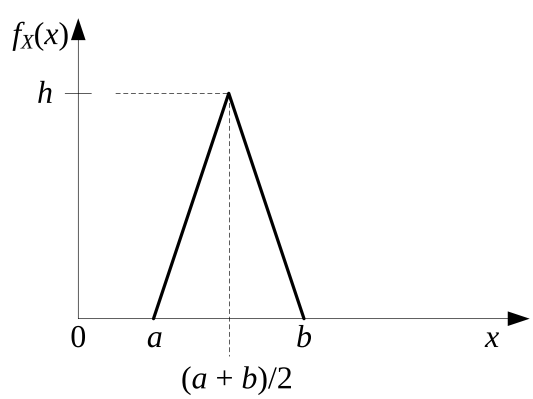
- za $p_1 \leq p_2 + p_3 + p_4$ vrijedi: $-p_1 + p_2 + p_3 + p_4 \leq p_1 + p_2 - p_3 - p_4$, tj.

$-p_1 + p_2 + 0{,}34 \leq p_1 + p_2 - 0{,}34$,

i konačno: $\boldsymbol{p_1 \geq 0{,}34}$

Kao što je već ranije rečeno, zbog jednakosti $p_1 + p_2 = 1 - (p_3 + p_4)$, slijedi $p_1 < 0{,}66 - 0{,}19$, tj. $\boldsymbol{p_1 < 0{,}47}$. Ova dva uvjeta za je moguće istovremeno zadovoljiti pa je konačno rješenje: $p_1 \in [0{,}34, 0{,}47)$.

Zadatak 8. (20 bodova) Slučajna varijabla X ima funkciju gustoće vjerojatnosti zadanu slikom. Odredite entropiju slučajne varijable X ako je zadano $a = 1$ i $b = 2$.



a) –0,193

b) 1,121

c) 0,548

d) –0,271

e) ništa od navedenog.

Postupak rješavanja:

Izraz za $f_X(x)$ određuje se pomoću jednadžbi pravaca kroz dvije točke. Funkcija $f_X(x)$ određena je sljedećim izrazom:

$$f_X(x)=\begin{cases}\dfrac{2h}{b-a}(x-a) & \text{za } a\le x\le \dfrac{a+b}{2}\\ \dfrac{2h}{b-a}(b-x) & \text{za } \dfrac{a+b}{2}\le x\le b\end{cases}$$

Entropiju, tj. diferencijalnu entropiju slučajne varijable X računamo prema sljedećem izrazu:

$$H(X)=-\int_a^{(a+b)/2}\frac{2h}{b-a}(x-a)\ln\left[\frac{2h}{b-a}(x-a)\right]dx-\int_{(a+b)/2}^{b}\frac{2h}{b-a}(b-x)\ln\left[\frac{2h}{b-a}(b-x)\right]dx$$

Koristeći identitet

$$\int x\ln(x)dx=\frac{x^2}{2}\ln(x)-\frac{x^2}{4}$$

dobivamo:

$$H(X)=\frac{-h}{b-a}\left\{(x-a)^2\ln\left[\frac{2h}{b-a}(x-a)\right]-\frac{(x-a)^2}{2}\right\}\Bigg|_a^{(a+b)/2}+$$

$$+\frac{h}{b-a}\left\{(b-x)^2\ln\left[\frac{2h}{b-a}(b-x)\right]-\frac{(b-x)^2}{2}\right\}\Bigg|_{(a+b)/2}^{b}$$

Nakon sređivanja gornjeg izraza dobivamo:

$$H(X)=\frac{h(b-a)}{2}\left[-\ln(h)+\frac{1}{2}\right]$$

S obzirom da mora vrijediti

$\int_{-\infty}^{\infty} f_X(x)dx=1$ jer je $f_X(x)$ funkcija gustoće vjerojatnosti slučajne varijable, u ovom konkretnom slučaju to implicira da je $h\cdot(b-a)/2 = 1$. Dakle, $H(X) = -\ln(h) + 1/2$. Budući da su zadani $a = 1$ i $b = 2$, mora vrijediti $h = 2$. Sukladno tome, $H(X) = -\ln(2) + 0,5 = -0,193$ (rješenje označeno slovom b).