Sveučilište u Zagrebu
Fakultet elektrotehnike i računarstva

Drugi ispitni rok (jesenski) iz predmeta **TEORIJA INFORMACIJE**, 2. rujna 2021.

Pravilo bodovanja zadataka

Svaki točno odgovoreni zadatak donosi 10 bodova, netočno odgovoreni 4 negativna boda, a neodgovoreni 0 bodova.

Zadatak 1. Zadan je skup simbola na izvoru komunikacijskog kanala, $X = \{x_1, x_2, ..., x_n\}$. Svakom simbolu pridijeljena je apriorna vjerojatnost pojavljivanja $P(x_i)$, $i = 1, ..., n$. Zbroj svih vjerojatnosti $P(x_i)$ iznosi 1. Nadalje, pretvorimo svaki simbol x_i u dva nova simbola, y_{2i-1} i y_{2i}, pri čemu za apriorne vjerojatnosti pojavljivanja simbola y_j, $j = 1, ..., 2n$, vrijedi sljedeće pravilo: $P(y_{2i-1}) = P(y_{2i}) = P(x_i)/2$, $\forall\ i = 1, ..., n$ Razlika između entropija ova dva skupa simbola, $H(X) - H(Y)$, iznosi

a) –2 bit/simbol;

b) –1 bit/simbol;

c) 2 bit/simbol;

d) 1 bit/simbol;

e) Ništa od navedenog.

Postupak rješavanja:

Entropija skupa simbola X dana je izrazom:

$$H(X) = -\sum_{i=1}^{n} P(x_i)\log_2\left[P(x_i)\right]\left[\text{bit/simbol}\right].$$

Entropija skupa simbola Y dana je izrazom:

$$H(Y) = -\sum_{j=1}^{2n} P(y_j)\log_2\left[P(y_j)\right] = -2\sum_{i=1}^{n}\frac{P(x_i)}{2}\log_2\left[\frac{P(x_i)}{2}\right] = -\sum_{i=1}^{n} P(x_i)\left\{\log_2\left[P(x_i)\right] - \log_2(2)\right\}$$

$$= -\sum_{i=1}^{n} P(x_i)\log_2\left[P(x_i)\right] + \log_2(2)\sum_{i=1}^{n} P(x_i) = H(X) + 1\left[\text{bit/simbol}\right]$$

Dakle, razlika entropija $H(X) - H(Y)$ iznosi –1 bit/simbol (rješenje označeno slovom b).

Zadatak 2. Abeceda izvora M sadrži 9 simbola, m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8 i m_9, s pripadajućim vjerojatnostima pojavljivanja simbola od $P(m_1)$ do $P(m_9)$ kako slijedi: 0,49, 0,14, 0,14, 0,07, 0,07. 0,04, 0,02, 0,02, 0,01. Koder informacije koristi tehniku Shannon-Fano i svaki simbol kodira određenim brojem binarnih simbola iz abecede {0, 1}. Odredite efikasnost kôda.

a) 0,938;

b) 0,975;

c) 0,993;

d) 0,959;

e) Ništa od navedenog.

Postupak rješavanja:

Ako provedemo kodiranje zadanog skupa simbola tehnikom Shannon-Fano, dobit ćemo sljedeće kodne riječi (kodiranje provedeno sukladno primjeru u udžbeniku "Uvod u teoriju informacije i kodiranje", 2. izdanje, Tablica 5.4, stranica 254):

simbol m_i	vjerojatnost $P(m_i)$	kodna riječ $\mathbf{c}(m_i)$	duljina l_i
m_1	0,49	0	1
m_2	0,14	100	3
m_3	0,14	101	3
m_4	0,07	1100	4
m_5	0,07	1101	4
m_6	0,04	1110	4
m_7	0,02	11110	5
m_8	0,02	111110	6
m_9	0,01	111111	6

Sukladno tome, prosječna duljina kodne riječi iznosi:

$$\overline{L} = \sum_{i=1}^{9} l_i P(m_i) = 2{,}33 \text{ bit/simbol}.$$

Entropija skupa simbola M iznosi:

$$H(M) = -\sum_{i=1}^{n} P(m_i) \log_2 \left[P(m_i) \right] = 2{,}314 \left[\text{bit/simbol} \right].$$

Konačno, efikasnost kôda iznosi:

$$\varepsilon = \frac{H(M)}{\overline{L}} = 0{,}993 \text{ (rješenje označeno slovom c).}$$

Zadatak 3. Binarni izvor generira dva simbola iz abecede $X = \{x_1, x_2\}$ s pripadajućim vjerojatnostima pojavljivanja $P(x_1) = 2/3$ i $P(x_2) = 1/3$. Nadalje, pretpostavimo da isti izvor kombinira simbole x_1 i x_2 u združene simbole abecede $Y = \{x_1x_1, x_1x_2, x_2x_1, x_2x_2\}$, $P(x_i,x_j) = P(x_i)\cdot P(x_j)$, $\forall\ i, j \in \{1, 2\}$. Odredite omjer efikasnosti kôda ako se Huffmanov kôd primijeni nad proširenom abecedom Y u odnosu na njegovu primjenu na početnu abecedu X.

a) 18/17;

b) 18/19;

c) 36/35;

d) 36/37;

e) Ništa od navedenog.

Postupak rješavanja:

Entropija skupa X određena je poznatim izrazom i označimo je kao $H_1(X)$, srednja duljina kodne riječi iznosi 1 bit/simbol (jer postoje samo dva simbola u abecedi i primjenjuje se Huffmanovo kodiranje). Dakle, efikasnost koda u prvom slučaju određena je kao

$$\varepsilon_1 = H_1(X) / \overline{L_1} = H_1(X)$$

Ako združujemo istovrsne simbole u novu abecedu Y, entropija se udvostručuje (dokaz jednostavan, prikazan na predavanjima). Ako provedemo Huffmanovo kodiranje nad abecedom Y, dobivamo sljedeći kod

x_1x_1	4/9	0
x_1x_2	2/9	10
x_2x_1	2/9	111
x_2x_2	1/9	110

Sam kod nije bitan, nego prosječna duljina kodne riječi koja iznosi 17/9 bit/simbol. Dakle,

$$\varepsilon_2 = H_2(X) / \overline{L_2} = 2H_1(X) / \overline{L_2}.$$

Konačno

$$\frac{\varepsilon_2}{\varepsilon_1} = \frac{2H_1(X)/\overline{L_2}}{H_1(X)} = \frac{2}{\overline{L_2}} = \frac{18}{17}\,\text{bit/simbol}.$$

Zadatak 4. 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 na sljedeći način:

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

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

a) 2,665 bit/simbol;

b) 0,685 bit/simbol;

c) 1,666 bit/simbol;

d) 0,809 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: $\sum_{i=1}^{5} P(x_i) = 0,25 + 0,4 + 0,15 + 0,15 + 0,05 = 1$

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

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

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

$P(y_4) = 0,1$

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:

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

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

Zadatak 5. Zadan je linearni binarni blok kôd K sa sljedećom matricom provjere pariteta **H**:

$$\mathbf{H}=\left[\mathbf{A}^{\mathrm{T}}|\mathbf{I}_3\right]=\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & a & 0 & 1 & 0 \\ 1 & 1 & b & 0 & 0 & 1 \end{bmatrix}, \quad a, b\in\left\{0,1\right\}.$$

Odredite binarne brojeve a i b tako da najmanja težina kodne riječi koda K, određenog gererirajućom matricom $\mathbf{G} = [\mathbf{I}_3|\mathbf{A}]$, a koja je različita od nule, iznosi 3.

a) a = 0, b = 0;

b) a = 0, b = 1;

c) a = 1, b = 0;

d) a = 1, b = 1;

e) Ništa od navedenog.

Postupak rješavanja:

Temeljem zadane matrice **H** možemo odredit matricu **G** koda K:

$$G=\left[\mathbf{I}_3|\mathbf{A}\right]=\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & a & b \end{bmatrix}.$$

Sada je moguće odrediti sve kodne riječi koda K, množenjem poruka od $\mathbf{d}_1$ = [0 0 0] do $\mathbf{d}_8$ = [1 1 1] s matricom G, tj. $\mathbf{c}_i = \mathbf{d}_i\mathbf{G}$, i = 1, …, 8:

$$K=\left\{\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & a & b \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & a & 1+b \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1+a & 1+b \\ 1 & 1 & 0 & 0 & 1 & 1+b \\ 1 & 1 & 1 & 0 & 1+a & b \end{matrix}\right..$$

Sukladno tome, težine kodnih riječi su sljedeće:

$w(\mathbf{c}_1) = 0$, $w(\mathbf{c}_2) = 1 + a + b$, $w(\mathbf{c}_3) = 3$, $w(\mathbf{c}_4) = 3 + a + (1\oplus b) \geq 3$, $w(\mathbf{c}_5) = 4$,

$w(\mathbf{c}_6) = 3 + (1\oplus a) + (1\oplus b) \geq 3$, $w(\mathbf{c}_7) = 3 + (1\oplus b) \geq 3$, $w(\mathbf{c}_8) = 3 + (1\oplus a) + b \geq 3$.

Napomena: u izrazima za težine kodne riječi znak + označava cjelobrojno zbrajanje (1 + 1 = 2), a znak ⊕ označava zbrajanje po modulu dva (1⊕1 = 0).

Dakle, promatranjem težine kodne riječi $\mathbf{c}_2$, kao potencijalnog kandidata za kodnu riječ najmanje težine koja je ujedno različita od nule, evidentno je da mora vrijediti $a = 1$ i $b = 1$. U bilo kojoj od preostale tri kombinacije ispada da najmanja težina kodne riječi koda K, a koja je različita od nule, iznosi manje od 3.

Zadatak 6. Informacijski izvor generira simbole iz skupa od 2048 simbola, $X = \{x_1, \ldots, x_{2048}\}$. Svi su simboli međusobno jednako vjerojatni i kodiraju se binarnim kodom Shannon-Fano. Svaka se poruka prije slanja AWGN kanalom kodira Hammingovim kodom Ham(r). Širina prijenosnog pojasa kanala iznosi 4 kHz, a omjer srednje snage signala prema srednjoj snazi šuma u kanalu iznosi 20 dB. Odredite koliko je simbola u sekundi moguće poslati takvim kanalom.

a) 1775,52 simbol/s;

b) 3624,45 simbol/s;

c) 2421,17 simbol/s;

d) 2657,93 simbol/s;

e) Ništa od navedenog.

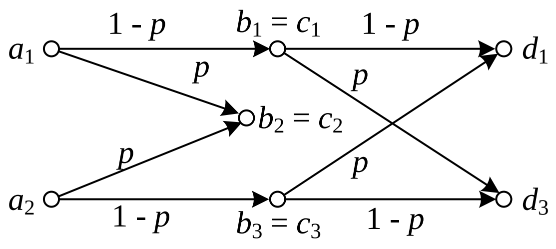
Postupak rješavanja:

Binarno kodiranje skupa od 2048 jednakovjerojatnih simbola tehnikom Shannon-Fano dat će u stvari ravnomjerni kod i svaki će simbol biti kodiran s 11 bita. Ako na takve poruke primijenimo Hammingov kod, bit će nužno na tih 11 bita nadodati 4 zaštitna bita i dobivamo kod Ham(4). Dakle, duljina svake kodne riječi iznosi 4 + 11 = 15 bita. Sam AWGN kanal ima kapacitet koji određujemo izrazom:

$$C = B\log_2\left(1+\frac{S}{N}\right)\left[\frac{\text{bit}}{\text{s}}\right].$$

Uvrštavanjem zadanih veličina dobivamo da je C = 26632,85 bit/s, a kad to podijelimo s brojem bita po simbolu, n = 15, konačni rezultat iznosi 1775,52 simbol/s.

Zadatak 7. Binarni kanal s brisanjem simbola vezna je u seriju s binarnim simetričnim kanalom, kako je prokazano na slici. Odredite vjerojatnost p tako da ukupni kanal ima obilježje slabo simetričnog kanala (WSC, engl. *weakly symmetric channel*).



a) 1/2;

b) 2/3;

c) 1/3;

d) 2/9;

e) Ništa od navedenog.

Postupak rješavanja:

Binarni kanal s brisanjem simbola ima matricu kanala određenu sljedećim izrazom:

$$\left[P\left(B|A\right)\right]=\begin{bmatrix}1-p & p & 0\\ 0 & p & 1-p\end{bmatrix}.$$

Da bi odredili matricu binarnog simetričnog kanala prikazanog na gornjoj slici, potrebno je uočiti da vrijedi da je $d_2 = b_2 = c_2$, tj. simbol koji je obrisan u binarnom kanalu s brisanjem simbola ostaje obrisan i u binarnom simetričnom kanalu. Sada možemo napisati modificiranu matricu binarnog simetričnog kanala:

$$\left[P\left(D|C\right)\right]=\begin{bmatrix}1-p & 0 & p\\ 0 & 1 & 0\\ p & 0 & 1-p\end{bmatrix}.$$

Za matricu ukupnog kanala vrijedi da je $[P(D|A)] = [P(B|A)]\cdot[P(D|C)]$, pa je ukupna matrica jednaka

$$\left[P\left(D|A\right)\right]=\left[P\left(B|A\right)\right]\cdot\left[P\left(D|C\right)\right]=\begin{bmatrix}1-p & p & 0\\ 0 & p & 1-p\end{bmatrix}\cdot\begin{bmatrix}1-p & 0 & p\\ 0 & 1 & 0\\ p & 0 & 1-p\end{bmatrix}=\begin{bmatrix}\left(1-p\right)^2 & p & \left(1-p\right)\cdot p\\ \left(1-p\right)\cdot p & p & \left(1-p\right)^2\end{bmatrix}.$$

Očito je da su oba retka permutacija jedan drugog, a da bi ukupni kanal imao obilježje slabo simetričnog kanala (WSC-a) zbroj svih stupaca mora biti međusobno jednak. Očito je da je zbroj elemenata prvog i trećeg stupca međusobno jednak pa preostaje da mora biti zadovoljena jednakost da je $(1-p)^2 + (1-p)\cdot p = 2p$. Nakon uređivanja lijeve strane jednakosti proizlazi da mora vrijediti da je $1-p = 2p$, tj. $p = 1/3$.

Zadatak 8. Promatrajte skup simbola X čiji su elementi simboli a i b, s vjerojatnostima pojavljivanja $P(a) = 1/3$ i $P(b) = 2/3$. Koristeći aritmetičko kodiranje izvor formira poruke duljine m simbola i šalje ih koderu kanala. Formirani aritmetički kôd je prefiksni, prema pravilu da se za svaku moguću poruku odredi njen kumulativni podinterval, uzme se vrijednost na sredini tog podintervala, zapiše se u binarnom obliku i od tog zapisa se uzme n znamenki (gledano desno od decimalne točke/zareza). Ako za barem jednu od 2^m poruka duljine m bita vrijedi $n \geq 10$, odredite koliko najmanje mora iznositi duljina poruke m.

a) 7 bit;

b) 6 bita;

c) 14 bita;

d) 15 bita;

e) Ništa od navedenog.

Postupak rješavanja:

Pravilo za formiranje prefiksnog aritmetičkog koda (vidi udžbenik na stranicama 264 i 265, stranicu 62 zbirke zadataka te slajdove s predavanja) je sljedeće:

Svakoj poruci duljine m bita pridružen je podinterval duljine $P(x)$ jednake umnošku vjerojatnosti svih simbola od kojih je poruka sastavljena;

Ako su poruke duljine m bita, onda se duljine njima odgovarajućih podintervala kreću u rasponu od $(1/3)^m$ do $(2/3)^m$;

Svakoj mogućoj poruci duljine m bita pripada kumulativni podinterval, a za svaki se kumulativni podinterval odabire broj na njegovoj sredini i pretvara u binarni ekvivalent;

Prefiksnost aritmetičkog koda postiže se tako da se od binarnog broja uzme broj znamenki $l(x)$ koji je s duljinom pripadajućeg podintervala $P(x)$ vezan izrazom:

$$l(x) = \left\lceil \log_2 \frac{1}{P(x)} \right\rceil + 1\,[\text{bit}];$$

Sukladno traženom da mora vrijediti $n \geq 10$, slijedi:

$$l(x) = \left\lceil \log_2 \frac{1}{P(x)} \right\rceil + 1 \geq 10;$$

Dakle, uvjet će najprije ispuniti poruka čija je duljina podintervala najmanja. Dakle, poruka sastavljena od m simbola a, $P(x) = (1/3)^m$, pa vrijedi:

$$l(x) = \left\lceil \log_2 \frac{1}{(1/3)^m} \right\rceil + 1 = \lceil m \cdot \log_2 3 \rceil + 1 \geq 10.$$

Očito je da za $m = 5$ uvjet nije ispunjen, dok je za $m = 6$ ispunjen pa prema tome najmanja duljina poruke m iznosi 6.

Zadatak 9. Neki ciklični kod K sadrži kodnu riječ 1 0 1 0 1 0 1. Broj riječi koda manji je od 128. Odredite duljinu poruke koja odgovara zadanoj kodnoj riječi.

a) 1 bit;

b) 6 bita;

c) 4 bita;

d) 3 bita;

e) Ništa od navedenog.

Postupak rješavanja:

Polinom kodne riječi je produkt polinoma poruke i generirajućeg polinoma. Do generirajućeg polinoma moguće je doći faktorizacijom polinom $x^7 + 1$:

$$x^7 + 1 = 1 \cdot (x+1) \cdot (x^3 + x + 1) \cdot (x^3 + x^2 + 1).$$

Ako bi odabrali polinom $g_1(x) = 1$, dobili bi sve riječi iz R_7, što znači da bi broj riječi u kodu iznosio 128, a zadano je da je broj riječi u kodu, M, manji od 128. Dakle, potencijalni polinomi su preostala tri iz rastava polinom $x^7 + 1$. Podijelimo li zadanu kodnu riječ $c(x)$ sa svakim od tih polinoma dobit ćemo sljedeće rezultate:

$$x^6+x^4+x^2+1 : x+1 = x^5+x^4+x+1$$

$$x^6+x^4+x^2+1 : x^3+x+1 = x^3+x \text{ i ostatak } x^2+x$$

$$x^6+x^4+x^2+1 : x^3+x^2+1 = x^3+x^2+1 \text{ i ostatak } x^2$$

$$x^6+x^4+x^2+1 : \left[(x+1)\left(x^3+x^2+1\right)\right] = x^6+x^4+x^2+1 : x^4+x^2+x+1 = x^2 \text{ i ostatak } x^3+1$$

$$x^6+x^4+x^2+1 : \left[(x+1)\left(x^3+x+1\right)\right] = x^6+x^4+x^2+1 : x^4+x^3+x^2+1 = x^2+x+1 \text{ i ostatak } x^2+x$$

$$x^6+x^4+x^2+1 : \left[\left(x^3+x^2+1\right)\left(x^3+x+1\right)\right] = x^6+x^4+x^2+1 : x^6+x^5+x^4+x^3+x^2+x+1 = 1 \text{ i ostatak } x^5+x^3+x$$

Dakle, kodna riječ je djeljiva bez ostatka jedinom s polinomom $x + 1$, što znači da je to generirajući polinom $g(x)$ koda K Sama poruka je, dakle, $c(x):g(x) = d(x) = 1\ 1\ 0\ 0\ 1\ 1$ i sadrži $k = 6$ bita.

Zadatak 10. U nekom kanalu u kontinuiranom vremenu omjer srednje snage signala prema srednjoj snazi šuma iznosi 10^5. Odredite koliko puta će se smanjiti prijenosna brzina u tom kanalu u odnosu na kapacitet kanala uslijed korištenja neoptimalnog kodnog sustava koji unosi smanjenje omjera srednje snage signala prema srednjoj snazi šuma u iznosu od 30 dB.

a) 3,149 puta;

b) 4,014 puta;

c) 1,667 puta;

d) 2,495 puta;

e) Ništa od navedenog.

Postupak rješavanja:

Zadan je omjer srednje snage signala prema srednjoj snazi šuma $S/N = 10^5$ te smanjenje omjera srednje snage signala prema srednjoj snazi šuma $\Gamma = 30$ dB, tj. $\Gamma = 1000$. Omjer kapaciteta kanala prema prijenosnoj brzini određen je izrazom:

$$\frac{C}{R} = \frac{B \cdot \log_2\left(1+\dfrac{S}{N}\right)}{B \cdot \log_2\left(1+\dfrac{S}{N\Gamma}\right)} = 2,495$$