Sveučilište u Zagrebu
Fakultet elektrotehnike i računarstva

3. ispitni rok (jesenski) iz predmeta **TEORIJA INFORMACIJE**, 7. rujna 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) Dan je skup simbola S s pripadajućim vjerojatnostima pojavljivanja :

$$S = \begin{pmatrix} s_1 & s_2 & & s_m \\ & & \cdots & \\ p_1 & p_2 & & p_m \end{pmatrix}$$

Simboli su jednoznačno kodirani prefiksnim kodom. Ako je $m = 6$ i ako su duljine kodnih riječi zadane izrazom $\{l_1, l_2, ..., l_6\} = \{1, 1, 2, 3, 2, 3\}$, odredite najmanji broj simbola abecede prefiksnog koda.

a) 4

b) 3

c) 2

d) 5

e) ništa od navedenog

Postupak rješavanja:

Za svaki prefiksni kod sa abecedom od d simbola i duljinama kodnih riječi l_1, l_2, ..., l_n vrijedi Kraftova nejednakost, $\sum_{i=1}^{m} d^{-l_i} \leq 1$, tj. za konkretan slučaj:

$$\frac{1}{d} + \frac{1}{d} + \frac{1}{d^2} + \frac{1}{d^3} + \frac{1}{d^2} + \frac{1}{d^3} \leq 1$$

$$\frac{2}{d} + \frac{2}{d^2} + \frac{2}{d^3} \leq 1$$

$$2d^2 + 2d + 2 - d^3 \leq 0$$

$$d \geq 2.9196$$

Kako broj simbola abecede mora biti prirodan broj, najmanji cijeli broj koji zadovoljava Kraftovu nejednakost je :

$$d_{\min} = \lceil 2.9196 \rceil = 3$$

$$d_{\min} = 3$$

Zadatak 2. (10 bodova) Razmatrajte blok kôd K s 8 kodnih riječi koji svaku poruku duljine 3 bita kodira dodatnim paritetnim bitom koristeći pri tome neparni paritet. Kodne riječi se prenose binarnim simetričnim kanalom u kojem vjerojatnost pogrešnog prijenosa bita iznosi 0,01. Odredite vjerojatnost da zadani kôd K otkrije pogreške bita, ako pretpostavimo da je na svakoj kodnoj riječi sigurno nastupila barem jedna pogreška. Napomena: pazite na uvjetnu vjerojatnost. Nerazumijevanje tog pojma nikoga ne amnestira pa neće biti prostora za "Ja sam to drugačije shvatio." Rješavati s oprezom!

a) $38,82 \cdot 10^{-3}$

b) $75,33 \cdot 10^{-3}$

c) 0,985

d) 0,965

e) ništa od navedenog

Postupak rješavanja:

U analizi prijenosa kodnih riječi promatranim binarnim simetričnim kanalom postoje sljedeći događaji: A – pogreška je otkrivena, A^* – pogreška nije otkrivena, B – pogreška je nastupila i B^* – pogreška nije nastupila. Vjerojatnost da je pogreška otkrivena ako je sigurno i nastupila na jednom ili više bita kodne riječi je uvjetna vjerojatnost $P(A|B)$ i vrijedi: $P(A|B) = P(A, B)/ P(B)$, pri čemu je $P(A, B)$ vjerojatnost da je pogreška otkrivena i da je nastupila te vrijedi:

$$P(A,B)=\underbrace{\binom{4}{1}p(1-p)^3}_{\text{jednostruka pogreška}}+\underbrace{\binom{4}{3}p^3(1-p)}_{\text{trostruka pogreška}}=38,82\cdot 10^{-3}.$$

Drugim riječima, pogreška će biti otkrivena ako je nastupila na jednom ili na tri bita kodne riječi. Vjerojatnost da je pogreška nastupila dana je izrazom:

$$P(B)=P(A,B)+P(A^*,B)\ ,$$

gdje je $P(A^*, B)$ vjerojatnost da je pogreška nastupila i nije otkrivena, što se događa ako pogreška nastupi na dva ili na četiri bita unutar kodne riječi:

$$P(A^*,B)=\underbrace{\binom{4}{2}p^2(1-p)^2}_{\text{dvostruka pogreška}}+\underbrace{\binom{4}{4}p^4}_{\text{četverostruka pogreška}}=0,59\cdot 10^{-3}.$$

Sukladno tome, $P(B)=1-(1-p)^4=39,41\cdot 10^{-3}$ pa je u konačnici tražena vjerojatnost

$$P(A|B)=\frac{P(A,B)}{P(B)}=0,985.$$

Zadatak 3. (10 bodova) Na ulaz AWGN-kanala dolazi slučajni signal $X(t)$ koji ima Gaussovu razdiobu amplituda i obilježje stacionarnog slučajnog procesa u širem smislu (WSS). Pretpostavimo da je njegovo očekivanje jednako nuli, a njegova standardna devijacija iznosi 1 V. Na taj signal aditivno djeluje bijeli Gaussov šum čije je očekivanje jednako nuli, a spektralna gustoća snage $S_N(f)$ jednaka 0,5 mW/Hz za $f \in \mathbf{R}$. Odredite iznos kapaciteta takvog kanala za slučaj kad širina prijenosnog pojasa kanala teži u beskonačnost.

a) 1443 nat/s

b) 1000 nat/s

c) 1000 bit/s

d) 693 nat/s

e) ništa od navedenog

Postupak rješavanja:

U slučaju kad širina prijenosnog pojasa teži u beskonačnost, izraz za kapacitet kanala je:

$$\lim_{B\to\infty} C = C_\infty = \frac{1}{\ln 2}\frac{S}{N_0}\left[\text{bit/s}\right] = \frac{S}{N_0}\left[\text{nat/s}\right].$$

Ako je očekivanje signala $X(t)$ jednako nuli, tada je njegova srednja snaga jednaka kvadratu standardne devijacije, tj. $S = \sigma_X^2 = 1$ W. Ako je spektralna gustoća snage $S_N(f)$ jednaka 0,5 mW/Hz za $f \in \mathbf{R}$, to znači da je $N_0/2 = 0,5$ mW/Hz, a sam N_0 iznosi 1 mW/Hz. U konačnici $C_\infty = 1000$ nat/s.

Zadatak 4. (10 bodova) Razmatrajte linearni binarni blok kôd K s oznakom $[n, k, 3]$ koji je ujedno i perfektan. Odredite koliko iznosi duljina kodne riječi koda K, ako vrijedi $\mathbf{G}\cdot\mathbf{H}^T = \mathbf{0}$, pri čemu je $\mathbf{G}$ generirajuća matrica koda K, a nul-matrica $\mathbf{0}$ ima 4 stupca.

a) 15 bita

b) 16 bita

c) 11 bita

d) 7 bita

e) ništa od navedenog

Postupak rješavanja:

Ako je kôd $[n, k, 3]$ perfektan, to znači da se sve kodne riječi iz $V(n)$ nalaze unutar kugli u čijim se središtima nalaze kodne riječi koda K. Tada vrijedi:

$$2^k = \frac{2^n}{\binom{n}{0}+\binom{n}{1}}.$$

Dakle, vrijedi da je $2^{n-k} = 1 + n$. Nadalje, s obzirom da matrica $\mathbf{G}$ ima dimenzije $k \times n$, a matrica $\mathbf{H}^T$ dimenzije $n \times n - k$, tada njihov produkt, tj. matrica $\mathbf{0}$, ima dimenzije $k \times n - k$. Ako je zadano da matrica $\mathbf{0}$ ima četiri stupca, to znači da je $n - k = 4$. Uvrštavanjem u prethodni izraz dobivamo da je $1 + n = 16$, tj. $n = 15$.

Zadatak 5. (10 bodova) Promatrajte kanal kojeg karakterizira svojstvo da su mu reci matrice kanala, $[P(Y|X)]$, permutacije jedan drugog, a zbroj članova matrice po svakom stupcu međusobno je jednak. Pri tome X predstavlja skup simbola na ulazu, a Y skup simbola na izlazu kanala. Matrica kanala zadana je sljedećim izrazom:

$$\left[P\left(Y|X\right)\right] = \begin{bmatrix} 1/3 & 1/3 & 1/6 & a \\ b & c & d & e \end{bmatrix}, 0 < a, b, c, d, e < 1$$

Odredite kapacitet kanala. Napomena: Permutacija brojeva q_1, q_2, q_3, q_4 (brojevi q_i predstavljaju prvi redak matrice kanala) je svaka uređena četvorka oblika (r_1, r_2, r_3, r_4) u kojoj se svaki od brojeva q_1, q_2, q_3, q_4 javlja točno jedanput, pri čemu brojevi r_i predstavljaju drugi redak matrice kanala.

a) 2 bit/simbol

b) 1,918 bit/simbol

c) 1,585 bit/simbol

d) 0,082 bit/simbol

e) ništa od navedenog

Postupak rješavanja:

S obzirom da zbroj elemenata po retku matrice kanala mora iznositi 1, slijedi da je a = 1/6. Pod uvjetom da je drugi redak permutacija prvog retka (dakle, sadrži dvije vjerojatnosti 1/3 i dvije vjerojatnosti 1/6) te uz zadani uvjet da je zbroj elemenata matrice kanala po svakom stupcu međusobno jednak, postoji samo jedno moguće rješenje, a to je:

$$P\left(Y|X\right)=\begin{bmatrix} 1/3 & 1/3 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/3 & 1/3 \end{bmatrix}$$

Očito se radi o djelomično simetričnom kanalu (engl. *weakly symmetric channel*) čiji se kapacitet računa prema izrazu:

$$C=\log\left[\mathrm{card}\left(Y\right)\right]-H\left(Y|x\right)$$

pri čemu je

$$H\left(Y|x\right)=\sum_{j=1}^{4}P\left(y_j|x_i\right)\log\left(\frac{1}{P\left(y_j|x_i\right)}\right),i\in\{1,2\}$$

Dakle, za proračun kapaciteta kanala dovoljno je izračunati entropiju $H(Y|x)$ za jedan redak matrice kanala. S obzirom da skup Y ima 4 člana, vrijedi log[card (Y)] = 2 te:

$$H\left(Y|x\right)=-2\frac{1}{3}\log_2\frac{1}{3}-2\frac{1}{6}\log_2\frac{1}{6}=\frac{2}{3}\log_2 3+\frac{1}{3}\log_2 6=\frac{1}{3}+\log_2 3$$

pa je kapacitet kanala jednak $C = 2 - 1/3 - \log_2(3) = 5/3 - \log_2(3) = 0{,}082$ bit/simbol.

Zadatak 6. (10 bodova) Neka slučajna varijabla X ima funkciju gustoće vjerojatnosti definiranu izrazom

$$f_X\left(x\right)=\begin{cases}\frac{1}{a}e^{-x/a}, & x\geq 0\\ 0, & x<0\end{cases},\ a>0$$

Odredite koliko mora iznositi konstanta a pa da diferencijalna entropija slučajne varijable X iznosi 1 bit/simbol.

a) $a = 2/e$

b) $a = 1$

c) $a = 1/e$

d) $a = \ln 2$

e) ništa od navedenog

Postupak rješavanja:

Diferencijalnu entropiju slučajne varijable X određujemo temeljem izraza:

$$H\left(X\right)=-\int_{-\infty}^{\infty}f_X\left(x\right)\log_2 f_X\left(x\right)dx\ \left[\mathrm{bit/simbol}\right]$$

Ako u taj izraz uvrstimo izraz za funkciju $f_X(x)$, nakon integracije dobit ćemo izraz:

$$H\left(X\right)=\log_2\left(ae\right)=\frac{1}{\ln 2}\ln(ae)$$

Kako bi diferencijalna entropija bila jednaka 1 bit/simbol, nužno je da vrijedi $ae = 2$ pa je evidentno da je $a = 2/e$.

Zadatak 7. (20 bodova) Neki izvor generira simbole iz abecede $X = \{x_1, ..., x_n\}$.

$$X = \begin{pmatrix} x_1 & \dots & x_n \\ p_1 = \dfrac{N_1}{M} & \dots & p_n = \dfrac{N_n}{M} \end{pmatrix}, \quad M = \sum_{i=1}^{n} N_i$$

Razdioba vjerojatnosti simbola određena je mjerenjem tijekom vremena na jako dugom nizu simbola (pretpostavka: $t \to \infty$). Izvor je pseudoslučajan, generira uzorak od M simbola i nakon toga taj uzorak ponavlja trajno. Izvor je također i ergodičan. Odredite S_i kao broj pojavljivanja simbola x_i po skupu, promatrano na jednom mjestu u nizu te po svim mogućim nizovima koje izvor može generirati. Temeljem općeg izraza odredite broj pojavljivanja simbola x_1 čija vjerojatnost iznosi 1/2 uz $M = 10$ te $n = 3$ i $p_2 = 1/5$. Napomena: u rješenju mora biti izveden i opći izraz za S_i.

a) 6300

b) 30240

c) 2520

d) 1260

e) ništa od navedenog.

Postupak rješavanja:

Ako je izvor ergodičan onda prosjek po skupu mora biti jednak prosjeku po vremenu. Broj različitih sljedova duljine M u kojima je razdioba vjerojatnosti jednaka zadanoj iznosi

$$B_S = \binom{M}{N_1}\cdot\binom{M-N_1}{N_2}\cdot\binom{M-N_1-N_2}{N_3}\cdots\binom{M-N_1-N_2-\ldots-N_{n-2}}{N_{n-1}}\cdot\binom{M-N_1-N_2-\ldots-N_{n-1}}{N_n}=$$

$$=\binom{M}{N_1}\cdot\binom{M-N_1}{N_2}\cdot\binom{M-N_1-N_2}{N_3}\cdots\binom{N_{n-1}+N_n}{N_{n-1}}\cdot\binom{N_n}{N_n}=$$

$$=\frac{M!}{N_1!\left(M-N_1\right)!}\cdot\frac{\left(M-N_1\right)!}{N_2!\left(M-N_1-N_2\right)!}\cdot\frac{\left(M-N_1-N_2\right)!}{N_3!\left(M-N_1-N_2-N_3\right)!}\cdots\frac{N_n!}{0!N_n!}=$$

$$=\frac{M!}{N_1!N_2!\cdots N_n!}$$

Za broj pojavljivanja simbola i na nekom mjestu u nizu, S_i, mora vrijediti: $N_i : M = S_i : B_s$

Dakle, $S_i = \dfrac{B_s N_i}{M} = \dfrac{M!}{N_1!N_2!\cdots N_n!}\cdot\dfrac{N_i}{M} = \dfrac{(M-1)!}{\prod_{i=1}^{n} N_i!}N_i$

Konačno, za $M = 10$, uz zadane vjerojatnosti i broj simbola $n = 3$, vrijedi $N_1 = 5$, $N_2 = 2$ i $N_3 = 3$. Uvrštavanjem tih brojeva dobivamo $S_1 = 1260$.

Zadatak 8. (20 bodova) Na ulazu diskretnog binarnog komunikacijskog kanala pojavljuju se dva simbola $X = \{x_1, x_2\}$ s pripadajućim vjerojatnostima pojavljivanja $p(x_1)$, odnosno $p(x_2)$. Odredite kapacitet danog kanala. Matrica uvjetnih vjerojatnosti prijelaza u kanalu zadana je kao

$$\left[p\left(y_j \mid x_i\right)\right]=\begin{bmatrix} 1 & 0 \\ 0,\ 5 & 0,\ 5 \end{bmatrix}$$

a) 1 bit/simbol

b) 0,6438 bit/simbol

c) 0,3219 bit/simbol

d) 0,1609 bit/simbol

e) ništa od navedenog

Postupak rješavanja:

$$p\left(x_1\right)=q$$

$$p\left(x_2\right)=1-q$$

$$\left[p\left(x_i, y_j\right)\right]=\left[p\left(x_i\right)\, p\left(y_j \mid x_i\right)\right]=\begin{bmatrix} q & 0 \\ 0,5\ (1-q) & 0,5\ (1-q) \end{bmatrix}$$

Zbrajanjem po stupcima matrice združenih vjerojatnosti dobivamo vjerojatnosti pojave simbola na izlazu kanala:

$$p\left(y_1\right)=0,5\ (1+q)$$

$$p\ \left(y_2\right)=0,5\ (1-q)$$

Zatim određujemo entropiju na izlazu kanala izrazom:

$$H(Y)=-\sum_{j=1}^{2} p\left(y_j\right) \log_2 p\left(y_j\right)$$

$$H(Y)=-\frac{(1+q)}{2}\ \log_2(1+q)-\frac{(1-q)}{2}\ \log_2(1-q)+1$$

Entropiju šuma računamo izrazom:

$$H(Y \mid X)=-\sum_{i=1}^{2} \sum_{j=1}^{2} p\left(x_i, y_j\right) \log_2\left(p\left(y_j \mid x_i\right)\right)$$

$$H(Y \mid X)=-\left[q \log_2 1+0,5\ (1-q) \log_2\ (0,5)+0,5(1-q) \log_2\ (1-q)\right]$$

$$H(Y \mid X)=1-q$$

Transinformaciju računamo izrazom:

$$I(X ; Y)=H(Y)-H(Y \mid X)$$

$$I(X ; Y)=-\frac{1}{2 \ln 2}\ \left((1+q)\ \ln(1+q)+(1-q)\ \ln(1-q)\right)+q$$

Kako bismo odredili vjerojatnost pojavljivanja ulaznog skupa simbola, moramo derivirati jednadžbu transinformacije i dobiveni izraz izjednačiti s nulom:

$$0 = -\frac{1}{2\ln 2}\left(\ln\left(1+q\right)-\ln\left(1-q\right)\right)+1$$

$$0 = -\ln\left(1+q\right)+\ln\left(1-q\right)+\ln\left(4\right)$$

$$0 = \ln\left(\frac{4\ \left(1-q\right)}{1+q}\right)$$

$$q = 0,6$$

Kapacitet zadanog kanala računamo izrazom

$$C = \max_{\{p(x_i)\}} I\left(X;Y\right) = \max_{\{p(x_i)\}}\ (H(Y)-H(Y\mid X))$$

$$C = 0,3219\,\frac{\text{bit}}{\text{simbol}}$$