Prva zadaća iz Teorije Informacije

ZADATAK 1.

Entropija [illegible]

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i) \; [\text{bit/simbol}]$$

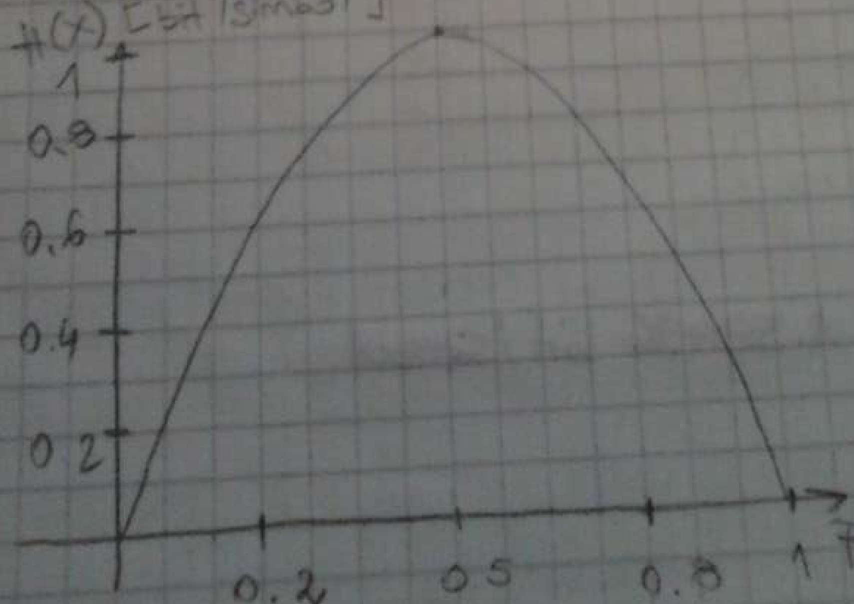
Uzajamni sadržaj informacije [illegible]

$$I(X;Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)}$$

a) $p \in [0,1]$

$$P(X=x) = \begin{cases} p & x = 1 \\ 1-p & x = 0 \end{cases}$$

$$H(X) = p \cdot \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}$$



[illegible]

$$H(X) = \frac{p \cdot \ln\left(\frac{1}{p}\right) + (1-p) \cdot \ln\left(\frac{1}{1-p}\right)}{\ln 2} = [illegible]$$

$$\frac{dH(X)}{dx} = [illegible]$$

$$= \frac{\ln\left(\frac{1-p}{p}\right)}{\ln 2}$$

$$\frac{dH(X)}{dx} = 0 \longrightarrow \frac{\ln\left(\frac{1-p}{p}\right)}{\ln 2} = 0$$

$$\frac{1-p}{p} = e^0$$

$$1-p = p \longrightarrow 2p = 1$$

$$p = 0.5$$

p je slučajna varijabla koja poprima vrijednost iz skupa $\{0, 0.5$ [illegible]

s jednakom vjerojatnošću

Koliko je $\mathbb{E}(H(X))$?

$$\mathbb{E}(H(X)) = \sum_{i=1}^{3} H(X : p = p_i) \cdot \frac{1}{3}$$

$$H(X : p = 0) = 0$$

$$H(X : p = 0.5) = 1$$

$$H(X : p = 1) = 0$$

$$\mathbb{E}(H(X)) = 0 \cdot \frac{1}{3} + \frac{1}{3} \cdot 1 + 0 \cdot \frac{1}{3} = \frac{1}{3} \text{ bit/simbol}$$

b) $X \sim \begin{pmatrix} -1 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$

$$H(X) = 3 \cdot \frac{1}{3} \cdot \log 3 = 1.58496 \text{ bit/simbol}$$

$X^2 \sim \begin{pmatrix} [illegible] \end{pmatrix} \sim \begin{pmatrix} [illegible] \end{pmatrix}$

$H(X^2)$ [illegible] $= 0.9183$ bit/simbol

$2^X \sim \begin{pmatrix} \frac{1}{2} & 1 & 2 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$ $H(2^X)$ [illegible] $= 1.58496$ bit/simbol

Obzirom na dobivene rezultate možemo zaključiti $H(V) \geq H(f(V))$

$H(V, f(V)) = H(V) + H(f(V)|V) = H(f(V)) + H(V|f(V))$

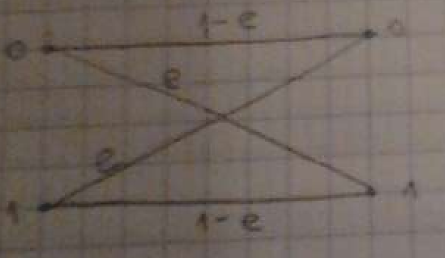
Obzirom da je $f(V)$ potpuno određena za svaku vrijednost V

vrijedi $H(f(V)|V) = 0$

$H(V|f(V)) \geq 0$ [illegible]

Iz gornjih svojstava dobivamo zaključak da je $H(V) \geq H(f(V))$

c)



$e = 1/8$

$X \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$ $X^2 \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$

$$[p(y_j|x_i)] = \begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & \frac{7}{8} & \frac{1}{8} \\ 1 & \frac{1}{8} & \frac{7}{8} \end{array}$$

$p(x_i) \cdot [p(y_j|x_i)] = p(x_i, y_j) = \begin{bmatrix} \frac{7}{24} & \\ \frac{1}{12} & \end{bmatrix}$

$H(X^2) = \frac{1}{3}\log 3 + \frac{2}{3}\log\frac{3}{2} = 0.9183$ bit/simb. $[p(y_j)] = [3/8 \quad 5/8]$

$H(Y) = \frac{3}{8}\log_2\frac{8}{3} + \frac{5}{8}\log_2\frac{8}{5} = 0.95443$ bit/simbol

$H(X^2, Y) = \frac{7}{12}\log_2\frac{12}{7} + \frac{7}{24}\log_2\frac{24}{7} + \frac{1}{12}\log_2 12 + \frac{1}{24}\log_2 24 = 1.4618$[illegible]

$I(X^2, Y) = H(X) + H(Y) - H(X, Y) = 0.41087$ bit/sim[illegible]

$$X^2 \sim \begin{pmatrix} 1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

$H(X^2) = \frac{1}{3}\log 3 + \frac{2}{3}\log\frac{3}{2} = 0.9183$ bit/simbol

$$2^X \sim \begin{pmatrix} \frac{1}{2} & 1 & 2 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \quad H(2^X) = 3 \cdot \frac{1}{3} \log 3 = 1.58496 \text{ bit/simbol}$$

Obzirom na dobivene rezultate možemo zaključiti $H(V) \geq H(f(V))$

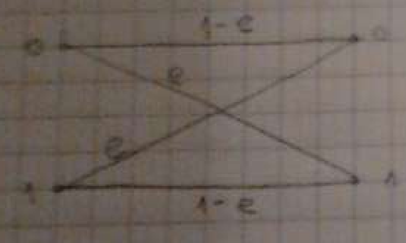
$H(V, f(V)) = H(V) + H(f(V)|V) = H(f(V)) + H(V|f(V))$

Obzirom da je $f(V)$ funkcija varijable V, ako znamo vrijednost V vrijedi $H(f(V)|V) = 0$

$H(V|f(V)) \geq 0$ ako i samo ako je funkcija [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]

Iz gornjih izraza [illegible] da je $H(V) \geq H(f(V))$

c)



$e = 1/8$

$$X \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \qquad X^2 \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

$$[p(y_j|x_i)] = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & \frac{7}{8} & \frac{1}{8} \\ 1 & \frac{1}{8} & \frac{7}{8} \end{array} \qquad p(x_i) \cdot [p(y_j|x_i)] = p(x_i, y_j) = \begin{bmatrix} \frac{7}{24} & \frac{1}{24} \\ \frac{1}{12} & \frac{7}{12} \end{bmatrix}$$

$H(X^2) = \frac{1}{3}\log 3 + \frac{2}{3}\log\frac{3}{2} = 0.9183$ bit/simb. $\quad [p(y_j)] = [3/8 \quad 5/8]$

$H(Y) = \frac{3}{8}\log_2\frac{8}{3} + \frac{5}{8}\log_2\frac{8}{5} = 0.95443$ bit/simbol

$H(X^2, Y) = \frac{7}{12}\log_2\frac{12}{7} + \frac{7}{24}\log_2\frac{24}{7} + \frac{1}{12}\log_2 12 + \frac{1}{24}\log_2 24 = 1.46186$ bit

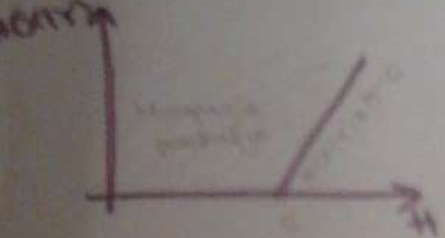
$$I(X^2; Y) = H(X) + H(Y) - H(X, Y) = 0.41087 \text{ bit/simbol}$$

Zadatak 2.

Kapacitet kanala je maksimalna transinformacija, gdje se maksimizacija provodi po svim mogućim razdiobama vjerojatnosti pojave simbola na ulazu.

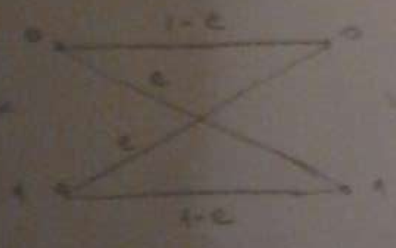
$$C = \max_{\{p(x_i)\}} I(X;Y) \ [\text{bit/simbol}]$$

Temeljni teorem kanala sa smetnjama.



Ako je $H \le C$, tada postoji takav sustav kodiranja da se poruke s izvora mogu prenositi kanalom sa proizvoljno malom učestalošću pogreške (tj. proizvoljno malom ekvivokacijom). Ako je $H > C$, tada je moguće kodirati tako da ekvivokacija bude manja od $H - C + \varepsilon$ [bit/simbol], gdje je ε proizvoljno mala veličina. Ne postoji metoda kodiranja kojom bi se postigla ekvivokacija manja od $H - C$ [bit/simbol].

a)



$$[p(y_j|x_i)] = \begin{bmatrix} 1-e & e \\ e & 1-e \end{bmatrix}$$

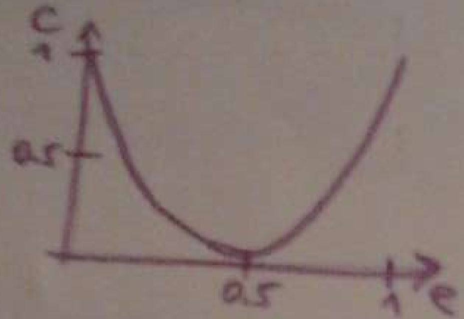
a) $$C = \max_{\{p(x_i)\}} I(X;Y) = \max_{\{p(x_i)\}} [H(Y) - H(Y|X)]$$

$H(Y)$ je maksimalna onda kada su vjerojatnosti pojave simbola jednake. Zbog simetričnosti kanala, pojave simbola na izlazu bit će simetrične ako su pojave simbola na ulazu simetrične. Dakle vrijedi $p(1) = p(0) = 1/2$

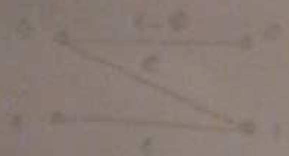
$$[p(x_i, y_j)] = [p(x_i)] \cdot [p(y_j|x_i)] = \begin{bmatrix} 0.5(1-e) & 0.5e \\ 0.5e & 0.5(1-e) \end{bmatrix}$$

$$p(y_j) = [0.5 \quad 0.5]$$

$C = 1 + e\log e + (1-e)\log(1-e)$ bit/simbol



b)



$$p(y_j|x_i) = \begin{array}{c} 0 \\ 1 \end{array} \begin{bmatrix} 1-e & e \\ 0 & 1 \end{bmatrix}$$

$p(x_i) = [\, p \quad 1-p \,]$

$$p(x_i, y_j) = \begin{bmatrix} p(1-e) & pe \\ 0 & 1-p \end{bmatrix} \qquad p(y_j) = [\, p(1-e) \quad pe + 1 - p \,]$$

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

Transinformacija biti će maksimalna onda kada je $H(Y)$ maksimalan a $H(Y|X)$ minimalan.

Iz grafa se vidi kako će kanal biti maksimalan ako vrijedi $e = 0$, budući da će tada $H(Y)$ biti maks

$p(x_i) = [\,p \quad 1-p\,]$

$p(y_j|x_i) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$H(Y) = -p\log p - (1-p)\log(1-p)$

$H(X) = -p\log p - (1-p)\log(1-p)$

$H(X,Y) = -p\log p - (1-p)\log(1-p)$

$I(X;Y) = -p\log p - (1-p)\log(1-p) - \dots$

$= -p\log p - (1-p)\log(1-p)$

$\frac{dI(X;Y)}{dp} = 0 = -\log p - 1 + \log(1-p) + 1$

$\log\frac{1-p}{p} = 0$

$\frac{1-p}{p} = 1$

$1-p = p$

$1 = 2p$

$p = 1/2$

$C = \max I(X;Y) = 1$ bit/simbol.

c)

Iz grafa je vidljivo kako će prijenos informacije biti minimal[an] kada je $e=1$. U tom slučaju koji god da se ulazni sim[bol] pojavi, izlazni će biti 1.

$p(y_j) = [\,0 \quad 1\,]$ $\quad p(y_j|x_i) = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$

$p(x_i) = [\,p \quad 1-p\,]$ $\quad p(x_i, y_j) = \begin{bmatrix} 0 & p \\ 0 & 1-p \end{bmatrix}$

$H(Y) = 0$

$H(X) = -p\log p - (1-p)\log(1-p)$

$H(X,Y) = -p\log p - (1-p)\log(1-p)$

$I(X;Y) = -p\log p - (1-p)\log(1-p) + p\log p + (1-p)\log(1-p)$

$C = 0$

p nije moguće jednoznačno odrediti!

a) [illegible] 1.2a, [illegible]

i) [illegible] 0.1 [illegible]

$$[p(y_j|x_i)] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$p(x_i) = 0.1$

$H(X) = 10 \cdot 0.1 \cdot \log 10$
$= \log_2 10$ bit/simbol

$$[p(x_i, y_j)] = [p(x_i)] \cdot [p(y_j|x_i)] = \begin{bmatrix} 0.1 & 0 & \cdots & 0 \\ 0 & 0.1 & & \\ 0 & 0 & 0.1 & \\ & & & 0.1 \end{bmatrix}$$

$p(y_j) = 0.1$

$H(Y) = \log_2 10$ bit/simbol

$C = \max I(X;Y) = \max (H(X) + H(Y) - H(X,Y))$

$H(X,Y) = 10 \cdot 0.1 \log 10 = \log 10$ bit/simbol

$C = \max I(X;Y) = \log 10 + \log 10 - \log 10 = \log_2 10$ bit/simbol

ii)

$$[p(y_j|x_i)] = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \end{bmatrix}$$

$p(x_i) = 0.1$

$H(X) = 10 \cdot 0.1 \log \frac{1}{0.1}$
$= \log_2 10$ bit/simbol

$$[p(x_i, y_j)] = [p(x_i)] \cdot [p(y_j|x_i)] = \begin{bmatrix} 0.05 & 0.05 & 0 & 0 & \cdots \\ \cdot & 0.05 & 0.05 & \cdots & \\ & & \ddots & & \\ 0.05 & & & & 0.05 \end{bmatrix}$$

$p(y_j) = 0.1$

$H(Y) = \log_2 10$

$H(X,Y) = 20 \cdot 0.05 \cdot \log \frac{1}{0.05} = \log_2 20 = \log_2 2 + \log_2 10 = 1 + \log_2 10$ bit

$C = \max I(X;Y) = \log_2 10 + \cancel{\log_2 10} - 1 - \cancel{\log_2 10} =$

$= \log_2 10 - 1$ bit/simbol

Zadatak 3.

Prosječna duljina kodne riječi, L, za neki kod definirana je kao suma svih duljina kodnih riječi za svaki simbol $l(x_i)$, pomnožene s vjerojatnošću pojavljivanja tog simbola, $p(x_i)$, ili zapisano:

$$L = \sum_{i=1}^{n} p(x_i)\, l(x_i) = \sum_{i=1}^{n} p_i l_i \text{ [bit/simbol]}$$

Prosječna duljina kodne riječi direktno daje očekivanu duljinu kodirane poruke.

Kraftova nejednakost: za svaki prefiksni kôd s abecedom od d simbola (baza kôda) i duljinama kodnih riječi $l_1, l_2, l_3, \dots, l_n$ vrijedi:

$$\sum_{i=1}^{n} d^{-l_i} \le 1$$

Optimalni kodovi su oni kodovi kod kojih je prosječna duljina riječi veća od entropije za najviše jedan bit:

$$H(X) \le L \le H(X) + 1$$

a) Shannon-Fanoovo kodiranje

					kod	duljina
a	0.5	0			0	1
b	0.125	1	0	0	100	3
c	0.125	1	0	1	101	3
d	0.125	1	1	0	110	3
e	0.125	1	1	1	111	3

$H(X) = 2,$

$H(X) = L \Rightarrow$ **kod je optimalan**

$$L = \sum_{i=1}^{n} p_i l_i = 0.5 + 4 \cdot \frac{3}{8} = 2$$

$$\sum_{i=1}^{n} d^{-l_i} \le 1$$

$$\sum_{i=1}^{n} 2^{-l_i} = \frac{1}{2} + \frac{1}{8} \cdot 4 = 1 \qquad \sum_{i=1}^{n} d^{-l_i} = 1 \checkmark$$

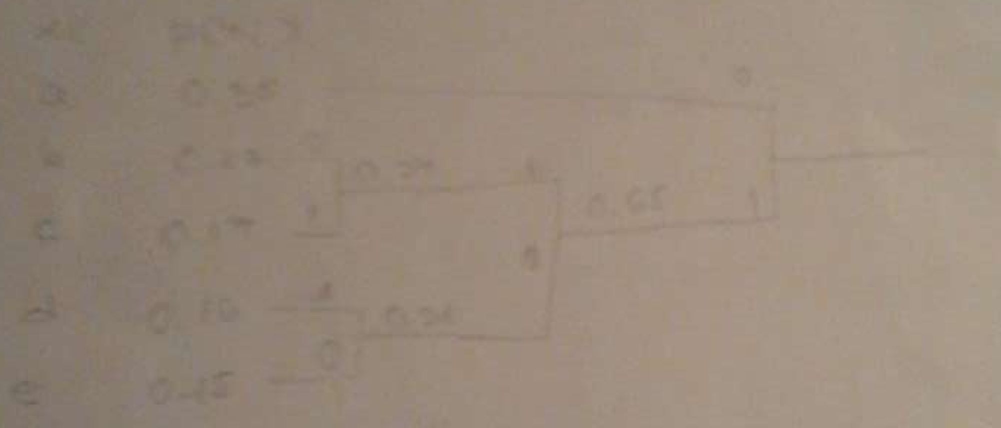
Primjer ne-optimalnog Shannon-Fanoovog koda:

[illegible]

$$\sum_{i=1}^{5} 2^{-l_i} \le 1$$

[illegible]

[illegible] Huffmanovim kodiranjem



x_i	$p(x_i)$	l_i
a	0	1
b	110	3
c	111	3
d	101	3
e	100	3

$$L = \sum_{i=1}^{5} p_i l_i = 2.3$$

Huffmanova metoda daje kod koji ima minimalnu prosječnu duljinu. Što pokazuje da Shannon-Fanovo kodiranje ne daje uvijek optimalna rješenja [illegible] koda

poruka ABA

$p(A) = 1/5$ $\qquad$ A: [0, 0.2)

$p(B) = 4/5$ $\qquad$ B: [0.2, 1)

donja granica = 0

A: $D' = D + (G-D)D_A = 0 + (1-0)\cdot 0 = 0$

$G' = D + (G-D)G_A = 0 + (1-0)\cdot 0.2 = 0.2$

$D = 0$
$G = 0.2$

B: $D' = D + (G-D)D_B = 0 + (0.2-0)\cdot 0.2 = 0.04$

$G' = D + (G-D)G_B = 0 + (0.2-0)\cdot 1 = 0.2$

$D = 0.04$
$G = 0.2$

A: $D' = D + (G-D)D_A = 0.04 + (0.2-0.04)\cdot 0 = 0.04$

$G' = D + (G-D)G_A = 0.04 + (0.2-0.04)\cdot 0.2 = 0.072$

Rješenje je u intervalu [0.04, 0.092)

Iz zadanog broja 0 [illegible]

0 pripada intervalu [0, 0.2) => **A** $\quad D = 0 \quad G = 0.2$

A: $D' = D + (G-D)D_A = 0 + (0.2-0)\cdot 0 = 0$

$G' = D + (G-D)G_A = 0 + (0.2-0)\cdot 0.2 = 0.04$

B: $D' = 0 + (0.2-0)\cdot 0.2 = 0.04$

$G' = 0 + (0.2-0)\cdot 1 = 0.2$

A $\rightarrow$ $D = 0$, $G = 0.04$

A: $D' = 0 + (0.04-0)\cdot 0 = 0$

$G' = 0 + (0.04-0)\cdot 0.2 = 8\cdot 10^{-3}$

B: $D' = 0 + (0.04-0)\cdot 0.2 = 8\cdot 10^{-3}$

$G' = 0 + (0.04-0)\cdot 1 = 0.04$

3. znak je **A**
$D = 0$
$G = 8\cdot 10^{-3}$

A: $D' = 0 + (8\cdot 10^{-3}-0)\cdot 0 = 0$

$G' = 0 + (8\cdot 10^{-3}-0)\cdot 0.2 = 1.6\cdot 10^{-3}$

B: $D' = 0 + (8\cdot 10^{-3}-0)\cdot 0.2 = 1.6\cdot 10^{-3}$

$G' = 0 + (8\cdot 10^{-3}-0)\cdot 1 = 8\cdot 10^{-3}$

4. znak je **A**
$D = 0$
$G = 1.6\cdot 10^{-3}$

[illegible]

Poruka je **AAAAA**

Za primljeni realni broj 0.21

0.21 spada u [0.2, 1)

1. znak je B

$D = 0.2$

$G = 1$

A: $D' = 0.2 + (1-0.2)\cdot 0 = 0.2$

$G' = 0.2 + (1-0.2)\cdot 0.2 = 0.36$

B: $D' = 0.2 + (1-0.2)\cdot 0.2 = 0.36$

$G' = 0.2 + (1-0.2)\cdot 1 = 1$

2. znak je A

$D = 0.2$

$G = 0.36$

A: $D' = 0.2 + (0.36-0.2)\cdot 0 = 0.2$

$G' = 0.2 + (0.36-0.2)\cdot 0.2 = 0.232$

B: $D' = 0.2 + (0.36-0.2)\cdot 0.2 = 0.232$

$G' = 0.2 + (0.36-0.2)\cdot 1 = 0.36$

3. znak je A

$D = 0.2$

$G = 0.232$

A: $D' = 0.2 + (0.232-0.2)\cdot 0 = 0.2$

$G' = 0.2 + (0.232-0.2)\cdot 0.2 = 0.2064$

B: $D' = 0.2 + (0.232-0.2)\cdot 0.2 = 0.2064$

$G' = 0.2 + (0.232-0.2)\cdot 1 = 0.232$

4. znak B

$D = 0.2064$

$G = 0.232$

znak je A

A: $D' = 0.2064 + (0.232 - 0.2064)\cdot 0 = 0.2064$

$G' = 0.2064 + (0.232 - 0.2064)\cdot 0.2 = 0.21152$

B: $D' = 0.2064 + (0.232 - 0.2064)\cdot 0.2 = 0.21152$

$G' = 0.2064 + (0.232 - 0.2064)\cdot 1 = 0.232$

Poruka je **BAABA**

77

[illegible] AACAACABCABAA*

Długie okna : 4

AACAACABCABAA*	(0, 0, A)
AACAACABCABAA*	(1, 1, C)
AACAACABCABAA*	(3, 3, A)
AACAACABCABAA*	(0, 0, B)
AACAACABCABAA*	(3, 3, A)
AACAACABCABAA*	(3, 1, *)