StuDocu.com

Examen 26 de marzo 2015, preguntas y respuestas

Informació i Seguretat (Universitat Autònoma de Barcelona)

StuDocu is not sponsored or endorsed by any college or university
Downloaded by Això L'audio (presidencialodisseu@gmail.com)

INFORMACIÓ I SEGURETAT
26 de març de 2015

Nom i cognoms: _____ Grup: _____

- Cal que justifiqueu convenientment totes les respostes
- Valoració dels exercicis: 1) 1.5+1 punts; 2) 1+1+0.5 punts; 3) 0.5+1+1 punts; 4) 1+0.75+0.75 punts
- $\log 3 = 1.58$, $\log 5 = 2.32$, $\log 7 = 2.8$

1. En un concurs de televisió els participants han de fer una sèrie d'eleccions per tal d'aconseguir el seu premi. Inicialment hi ha tres cofres: A, B i C. Cada cofre conté dos sobres amb premis. El cofre A conté un sobre amb 1500€ i un altre amb 1€. El cofre B, un sobre amb 5000€ i un amb 1000€. El cofre C, un amb 2€ i l'altre amb 0€. El participant ha d'escollir a l'atzar un cofre i després també a l'atzar un dels dos sobres del cofre.

   Considereu X l'elecció del cofre i Y el premi obtingut.

   (a) Calculeu $H(X)$. Digueu, segons la teoria vista a l'assignatura, quins són els valors màxims i mínims de $H(X|Y)$. Calculeu en aquest cas $H(X|Y)$ i justifiqueu el resultat obtingut.

   (b) Calculeu $H(Y)$, $H(Y|X)$. Digueu quina informació aporta el premi obtingut sobre el cofre escollit.

   Solució:

   (a) Definim $X = \{A, B, C\}$ i $Y = \{5000€, 1500\ €, 1000€, 2€, 1€, 0€\}$.
   Tenim que $p(A) = p(B) = p(C) = \frac{1}{3}$.
   Tenim que $H(X) = H(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) = \log 3 = 1.58$ bits. Sabem que $0 \leq H(X|Y) \leq H(X)$.
   Per calcular $H(X|Y)$, calculem les probabilitats condicionades $p(x_i|y_j)$, les conjuntes $p(x_i, y_j)$ i les condicionades $p(y_j|x_i)$.

$p(y_i\|x_j)$	5000	1500	1000	2	1	0
A	0	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0
B	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0	0
C	0	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$

$p(y_i, x_j)$	5000	1500	1000	2	1	0
A	0	$\frac{1}{6}$	0	0	$\frac{1}{6}$	0
B	$\frac{1}{6}$	0	$\frac{1}{6}$	0	0	0
C	0	0	0	$\frac{1}{6}$	0	$\frac{1}{6}$

$p(x_i\|y_j)$	5000	1500	1000	2	1	0
A	0	1	0	0	1	0
B	1	0	1	0	0	0
C	0	0	0	1	0	1

   Ara, $H(X|Y) = 6 \cdot \frac{1}{6} \log 1 + 12 \cdot 0 \log 0 = 0$. Com $H(X|Y) = 0$, alehores vol dir que si coneixem Y aleshores X queda completament determinat; és a dir, si coneixem el premi, sabem quin era el cofre escollit.

   (b) De les taules de l'apartat anterior podem extreure que $p(Y = y_j) = \frac{1}{6}$ per $j = 1, \ldots, 6$.
   Per tant, $H(Y) = H(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}) = \log 6 = \log 3 + \log 2 = 2.58$ bits.
   També de les taules podem extreure que $H(Y|X) = 6 \cdot \frac{1}{6} \log 2 + 12 \cdot 0 \log 0 = 1$ bit.
   Finalment, la informació que aporta el premi obtingut sobre el cofre escollit és $I(X, Y) = H(Y) - H(Y|X) = 2.58 - 1 = 1.58$ bits. També es podria calcular $I(X, Y) = H(X) - H(X|Y) = \log 3 - 0 = \log 3$.

Downloaded by Això L'audio (presidencialodisseu@gmail.com)

2. Una cadena de caràcters m s'ha comprimit utilitzant l'algorisme LZ77 obtenint la codificació $(0,0,\texttt{a})(0,0,\texttt{b})(0,0,\texttt{r})(3,1,\texttt{c})(2,1,\texttt{d})(7,4,\texttt{d})$.

   (a) Descomprimiu el missatge i recupereu la cadena m.

   (b) Doneu la codificació del mateix missatge m utilitzant l'algorisme LZ78.

   (c) Calculeu els percentatges de compressió de cada codificació suposant que els caràcters es codifiquen en 8 bits i les posicions en 4 bits

   Solució:

   (a) Descodificant amb LZ77 s'obté el missatge $m =$`'abracadabrad'`.

   (b) Si apliquem l'algorisme LZ78 obtenim,

	Dicc	*Codi*		*Dicc*	*Codi*
0	`null`				
1	`a`	`(0,a)`	5	`ad`	`(1,d)`
2	`b`	`(0,b)`	6	`ab`	`(1,b)`
3	`r`	`(0,r)`	7	`ra`	`(3,a)`
4	`ac`	`(1,c)`	8	`d`	`(0,d)`

   (c) En primer lloc, $|m| = 12 \times 8 = 96$ bits. La compressió LZ77 té $6 \times (4+4+8) = 96$ bits. Per tant, el percentatge de compressió serà $(1 - \frac{96}{96})\% = 0\%$. La compressió LZ78 té $8 \times (4+8) = 96$ bits. El percentatge de compressió també serà 0%.

3. "Quina és la contrasenya, Dr. Watson?" demanà Sherlock Holmes. "Es tracta d'una contrasenya que s'ha comprimit fent servir un codi binari. El resultat de la compressió és 000101101111101. A més tenim una taula amb 3 possibles codis, però no sabem quin és" respongué el seu company. El Dr. Watson li va ensenyar la taula següent:

$Missatge$	C_1	C_2	C_3
a_1	000	00	00
a_2	001	01	01
a_3	010	11	100
a_4	011	101	101
a_5	100	0111	1100
a_6	101	1011	1101
a_7	110	01000	1110
a_8	111	01001	1111

   "Molt interessant, Dr. Watson. De fet no cal considerar els 3, només hem de considerar aquells que siguin de descodificació única." El Dr. Watson va assentir. "Tot i així, tenim més d'un codi. Sort que hem trobat una altra nota on diu que el codi que s'ha fet servir és òptim i que les probabilitats dels símbols són $p(a_1) = \frac{3}{16}, p(a_2) = p(a_3) = \cdots = p(a_7) = \frac{2}{16}, p(a_8) = \frac{1}{16}$. Ara només hem de calcular l'eficiència dels codis i escollir el que tingui eficiència 1".

   (a) Justifiqueu quins són els codis de descodificació única i quins no.

   (b) Doneu la fòrmula de l'eficiència d'un codi binari i justifiqueu en quins casos l'eficiència d'un codi òptim és 1. En aquest cas tindrà raó el Dr. Watson i l'eficiència del codi òptim és 1?

Downloaded by Això L'audio (presidencialodisseu@gmail.com)

(c) Doneu la longitud mitjana dels codis de descodificació única i determineu quin és el codi que s'ha fet servir per comprimir la contrasenya. Demostreu que efectivament és òptim. Quina és la contrasenya?

Solució:

(a) Tenim que C_1 i C_3 són codis instantanis i, per tant, de descodificació única. El codi C_2 presenta ambigüitats; per exemple, 0111 podria ser a_2a_3 o a_5. Per tant C_2 no és de descodificació única.

(b) L'eficiència d'un codi binari és $\eta = \frac{H(S)}{\bar{L}}$. L'eficiència és 1 si i només si les probabilitats són de la forma 2^{-L_i}. En aquest cas, l'eficiència no serà 1 ja que $p(a_1)$ no és una potència de 2.
Una altra manera de veure-ho és la següent. Diem que l'eficiència és 1 si $H(S) = \bar{L}$. $H(S) = \frac{3}{16}\log\frac{16}{3} + 6\cdot\frac{2}{16}\log 8 + \frac{1}{16}\log 16 = \frac{52-3\cdot\log 3}{16} = \frac{47}{16} = 2.96$. Com les longituds mitjanes de C_1 i C_3 són $\frac{48}{16} = 3$ i $\frac{50}{16} = 3.125$ respectivament, tenim que en cap cas coincideix amb el valor de l'entropia i, per tant, l'eficiència no pot ser 1.

(c) Les longituds mitjanes de C_1 i C_3 són $\frac{48}{16} = 3$ i $\frac{50}{16} = 3.125$ respectivament. El codi que s'ha fet servir és C_1. Per determinar que efectivament és òptim, primer construïm un codi òptim fent servir l'algoritme de Huffman:

$$C_H = \{111, 110, 101, 100, 011, 010, 001, 000\}.$$

La longitud mitjana de C_H és 3 que coincideix amb la longitud mitjana de C_1; per tant, C_1 és òptim. La contrasenya, fent servir el codi C_1 és $a_1a_6a_6a_8a_6$.

4. Sigui $\{A_1, A_2, A_3, A_4\}$ el conjunt d'entrades i $\{B_1, B_2, B_3, B_4\}$ el de sortides d'un canal discret i sense memòria, amb matriu de probabilitats condicionades:

$$\begin{pmatrix} 1/2 & 1/4 & 1/4 & 0 \\ 1/4 & 1/2 & 0 & 1/4 \\ 0 & 1/4 & 1/4 & 1/2 \\ 1/4 & 0 & 1/2 & 1/4 \end{pmatrix}.$$

(a) Quina seria la probabilitat mitjana d'error descodificant a màxima versemblança?

(b) Quina és la capacitat del canal i la distribució inicial que fa que s'assoleixi la capacitat?

(c) Doneu, si existeix, una distribució de probabilitats inicial que faci que la informació mútua entre l'entrada i la sortida del canal sigui 2.

Solució:

(a) Fixant-nos en els valors màxims a cada columna obtenim la funció de descodificació a màxima versemblança:

$$\begin{aligned} B_1 &\longrightarrow A_1 \\ B_2 &\longrightarrow A_2 \\ B_3 &\longrightarrow A_4 \\ B_4 &\longrightarrow A_3 \end{aligned}$$

Diguem $p_i = P(A_i)$, per a $i = 1, \ldots, 4$, on $p_1 + p_2 + p_3 + p_4 = 1$. Aleshores, la probabilitat mitjana d'error en la descodificació és:

$$\overline{P}_e = 1 - p_1\frac{1}{2} - p_2\frac{1}{2} - p_3\frac{1}{2} - p_4\frac{1}{2} = 1 - (p_1 + p_2 + p_3 + p_4)\frac{1}{2} = \frac{1}{2}.$$

Downloaded by Això L'audio (presidencialodisseu@gmail.com)

(b) La capacitat val:

$$C = \log_2 4 - H(1/2, 1/4, 1/4) = 2 - 1.5 = 0.5 \text{ bits/entrada.}$$

Aquest màxim s'assoleix quan la distribució inicial és l'equiprobable.

(c) El valor màxim de la informació mútua entre l'entrada i la sortida del canal és la capacitat del canal. Per l'apartat anterior, aquest valor és 0.5 i, per tant, la informació mútua entre l'entrada i la sortida no pot ser 2.

Downloaded by Això L'audio (presidencialodisseu@gmail.com)