Examen 12 de abril 2012, preguntas y respuestas

Informació i Seguretat (Universitat Autònoma de Barcelona)

INFORMACIÓ I SEGURETAT
12 d'abril de 2012

Nom i cognoms: ______________________________ Grup: ______

- Cal que justifiqueu convenientment totes les respostes
- $\log 3 = 1.5849$, $\log 5 = 2.3219$, $\log 7 = 2.8073$

1. (25%) Volem endevinar un número entre 1 i 100. Suposant que cada número pot ser endevinat amb igual probabilitat,

   (a) Descriviu l'experiment com una font d'informació, és a dir, quin seria el conjunt S i les probabilitats associades.

   (b) Quina incertesa tenim a priori sobre el número que volem endevinar?

   (c) Si a les preguntes que fem només podem respondre `SI`, `NO`, quin és el nombre de preguntes que hauríem de fer per endevinar el número?

   Solució:

   (a) El nostre experiment està format per $S = \{1, \ldots, 100\}$ tots equiprobables, $p(i) = \frac{1}{100} =$.

   (b) La incertesa és igual a la entropia de la font, $H(S) = \log 100 = 6.6438$ bits.

   (c) Cada resposta ens aporta $\log 2 = 1$ bits d'informació. Si k és el nombre mínim de respostes, aleshores $k \log 2 \geq \log 100$ i, per tant, $k \geq 6.6438$. Necessitarem 7 preguntes.

2. (25%) Considereu una font d'informació formada per 5 símbols $S = \{a_1, a_2, a_3, a_4, a_5\}$ tots equiprobables.

   (a) Es possible construir un codi binari instantani per aquesta font de longituds $L_1 = L_2 = L_3 = L_4 = L_5 = 2$?

   (b) I de longituds $L_1 = L_2 = L_3 = 2$ i $L_4 = L_5 = 3$?

   (c) La font S l'hem codificada amb el codi $C_1 = \{10, 01, 00, 11, 110\}$. Calculeu l'eficiència d'aquest codi i justifiqueu si aquest codi és òptim.

   (d) Calculeu un codi binari instantani òptim per aquesta font i la seva eficiència.

   Solució:

   (a) No és possible ja que amb 2 bits només podem codificar 4 missatges.

   (b) Com que $3 \cdot 2^{-2} + 2 \cdot 2^{-3} = 1$, per la desigualtat de Kraft podem afirmar que podem construir un codi binari instantani amb aquestes longituds.

   (c) L'entropia de la font és $H(S) = \log 5 = 2.3219$. La longitud mitjana del codi serà $\bar{L} = 2.2$. L'eficiència serà $\eta = \frac{\log 5}{2.2} = 1.055$. L'eficiència no pot ser més gran de 1. Per tant, aquest codi no pot ser instantani ni tampoc òptim.

   (d) Aplicant l'algorisme de Huffman obtenim el codi $C = \{10, 111, 110, 01, 00\}$ amb una longitud mitjana $\bar{L} = 2.4$, $\eta = 0.9674$.

3. (25%) Considereu el text `m=''ABBCBCABABCAABCAAB''`.

(a) Comprimiu el text `m` utilitzant l'algorisme `LZ77` on tant la mida del diccionari com del buffer es representa amb 2 bits.

(b) Si cada caràcter es codifica en 8 bits, calculeu la taxa de compressió i el percentatge de compressió del missatge comprimit.

(c) Descomprimiu el missatge següent que també ha estat comprimit amb l'algorisme `LZ77`: `(0,0,A),(0,0,B),(0,0,R),(3,1,C),(2,1,D),(2,1,B),(0,0,R),(0,0,A)`.

Solució:

(a) La compressió seria: `(0,0,A),(0,0,B),(1,1,C),(2,2,A),(3,1,A),(2,1,C),(3,1,A),(0,0,B),(0,0,C)(3,1,A)(0,0,B)`.

(b) El missatge original té una mida, $18 \times 8 = 144$ bits. El missatge coprimit tindrà $11 \cdot (2+2+8) = 132$ bits. $R = \frac{132}{144} = 0.92$ bpb i el percentange de compressió serà 8%.

(c) El missatge descomprimit és `''ABRACADABRA''`.

4. (25%) Considereu el canal determinat per la matriu de transicions,

$$\Pi = \begin{pmatrix} 1/3 & 2/3 & 0 \\ 0 & 1/3 & 2/3 \\ 2/3 & 0 & 1/3 \end{pmatrix}$$

(a) Digueu de quin tipus de canal es tracta i quina és la capacitat del canal.

(b) Doneu una distribució inicial i final tal que assoleixi la capacitat del canal.

(c) Doneu la informació mútua de l'entrada i la sortida si la distribució inicial de probabilitats és $\{1/2, 1/4, 1/4\}$.

Solució:

(a) Es un canal totalment simètric amb $C = \log m - H = \log 3 - H(2/3, 1/3) = \log 3 - 0.9183 = 0.67$ bits.

(b) La capacitat s'assoleix quan les entrades són equiprobables, $p_1 = p_2 = p_3 = \frac{1}{3}$. I, en aquest cas, les sortides també seran equiprobables.

(c) La situació inicial serà:

$p(a)$	$p(b\|a)$		
1/2	1/3	2/3	0
1/4	0	1/3	2/3
1/4	2/3	0	1/3

Calculem les probabilitats conjuntes i les probabilitats dels símbols de sortida:

	$p(a,b)$		
	1/6	2/6	0
	0	1/12	2/12
	2/12	0	1/12
$p(b)$	4/12	5/12	3/12

$I(A,B) = H(B) - H(B|A) = 1.5546 - 0.9183 = 0.6363$ bits ja que,
$H(B) = H(4/12, 5/12, /3/12) = 1.5546$ i $H(B|A) = \sum p_i H = H = 0.9183$.