Examen 27 de marzo 2014, preguntas y respuestas

Informació i Seguretat (Universitat Autònoma de Barcelona)

INFORMACIÓ I SEGURETAT
27 de març de 2014

Nom i cognoms: ____ Niu: _______

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

1. Considereu el següent experiment aleatori: llencem una moneda simètrica. Si surt cara, llavors es tria un nombre aleatori entre 1 i 4 (tots quatre són igualment probables). Si surt creu, triem un nombre aleatori entre 2 i 5. Calculeu quina informació ens aporta saber el nombre escollit respecte del resultat de llençar la moneda (informació mútua).

Solució:

Considerem els esdeveniments:
$A_1 =$ "cara"; $A_2 =$ "creu"; $B_i =$ "es tria i" $(1 \leq i \leq 5)$.

Aleshores, hem de calcular la informació mútua entre els conjunts:

$$A = \{A_1, A_2\} \quad \text{i} \quad B = \{B_1, B_2, B_3, B_4, B_5\}.$$

Les probabilitats condicionades i conjuntes són:

$P(B_j \mid A_i)$	B_1	B_2	B_3	B_4	B_5	$P(A_i, B_j)$	B_1	B_2	B_3	B_4	B_5
A_1	1/4	1/4	1/4	1/4	0	A_1	1/8	1/8	1/8	1/8	0
A_2	0	1/4	1/4	1/4	1/4	A_2	0	1/8	1/8	1/8	1/8

Sumant les columnes de la taula de probabilitats conjuntes obtenim la distribució de B: $\{1/8, 1/4, 1/4, 1/4, 1/8\}$. Aleshores:

$$H(B) = 2\frac{1}{8}\log 8 + 3\frac{1}{4}\log 4 = \frac{3}{4} + \frac{6}{4} = \frac{9}{4} \quad \text{bits/resultat.}$$

$$\begin{aligned} H(B \mid A) &= \frac{1}{2}H(1/4, 1/4, 1/4, 1/4, 0) + \frac{1}{2}H(0, 1/4, 1/4, 1/4, 1/4) \\ &= H(1/4, 1/4, 1/4, 1/4) = \log 4 = 2 \quad \text{bits/resultat.} \end{aligned}$$

Per tant, la informació mútua és:

$$I(A, B) = H(B) - H(B \mid A) = \frac{9}{4} - 2 = \frac{1}{4} = 0.25 \quad \text{bits/resultat.}$$

2. Donada la font sense memòria $S = \{a_1, \ldots, a_5\}$, amb distribució de probabilitats: $p_1 = 1/3$, $p_2 = p_3 = 1/4$, $p_4 = p_5 = 1/12$,

   (a) Existeix algun codi a descodificació única amb eficiència igual a 1?

   (b) Existeix algun codi binari a descodificació única amb paraules-codi de longituds 2, 2, 2, 2 i 3?

   (c) Construïu un codi binari òptim i calculeu la seva eficiència.

   Solució:

   (a) Aplicant el primer teorema de Shannon tenim que

$$\eta = 1 \iff p_i = D^{-L_i} \;\; \forall i = 1, \ldots, 5.$$

   Però en aquest cas no existeix cap D que verifiqui això (per exemple, p_1 és potència de 3 i cap altra probabilitat és potència de 3). Per tant, no existeix cap codi a descodificació única amb eficiència $\eta = 1$.

   (b) Mirem si es verifica la desigualtat de McMillan:

$$4 \cdot 2^{-2} + 2^{-3} = 1 + \frac{1}{8} > 1.$$

   Per tant, no es verifica. O sigui, que no existeix un tal codi.

   (c) Aplicant el mètode de Huffman obtenim $\{00, 01, 10, 110, 111\}$ (hi pot haver canvis de zeros per uns però, en aquest cas, les longituds han de ser necessàriament aquestes). Calculem l'entropia:

$$\begin{aligned} H(S) &= \frac{1}{3}\log 3 + 2\frac{1}{4}\log 4 + 2\frac{1}{12}\log 12 = \frac{1}{3}\log 3 + 1 + \frac{1}{6}(\log 3 + \log 4) \\ &= \frac{1}{2}\log 3 + \frac{4}{3} = \frac{3\log 3 + 8}{6} \quad \text{bits/missatge.} \end{aligned}$$

   La longitud mitjana és:

$$\overline{L} = 2\frac{1}{3} + 2\frac{1}{4} + 2\frac{1}{4} + 3\frac{1}{12} + 3\frac{1}{12} = \frac{2}{3} + 1 + \frac{1}{2} = \frac{13}{6}.$$

   Finalment, l'eficiència és:

$$\eta = \frac{H(S)}{\overline{L}\log D} = \frac{\frac{3\log 3+8}{6}}{\frac{13}{6} \cdot \log 2} = \frac{3\log 3 + 8}{13} \approx \frac{12.74}{13} = 0.98.$$

3. (a) La següent seqüència és el resultat d'aplicar l'algorisme LZ77 a un text. Descobriu quin és el text.

$$(0,0,A), (0,0,S), (0,0,E), (0,0,R), (2,1,J), (2,1,\_\,), (3,1,A),$$
$$(3,1,D), (8,4,D), (5,3,B), (2,1,T), (0,0,U)(8,6,.)$$

(b) Codifiqueu, utilitzant l'algorisme LZ78 el text: TARARA_TARARA.

Solució:

(a) El text resultant és: ASEREJE_JA_DEJE_DEJEBETUDEJEBE.

(b) S'obté la seqüència de parelles: (0,T),(0,A),(0,R),(2,R),(2,_),(1,A),(3,A),(7,.).

4. Tenim un canal discret i sense memòria amb tres entrades $\{A_1, A_2, A_3\}$, tres sortides $\{B_1, B_2, B_3\}$ i matriu de probabilitats condicionades:

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

(a) Digueu quant val la capacitat d'aquest canal i les distribucions inicial i final que maximitzen la informació mútua.

(b) Quant val la probabilitat mitjana d'error si es descodifica a màxima versemblança?

(c) Si la distribució inicial fos $P(A_1) = 1/2$, $P(A_2) = 1/3$, $P(A_3) = 1/6$; digueu quant val la probabilitat mitjana d'error si es descodifica a mínima probabilitat d'error.

(a) És un canal simètric, per tant, la capacitat val

$$C = \log m - H = \log 3 - H(1/2, 1/2) \approx 1.58 - 1 = 0.58 \quad \text{bits/entrada}.$$

Aquest màxim de la informació mútua s'assoleix quan la distribució final és equiprobable. En tal cas, una distribució inicial que sempre fa que la final sigui equiprobable, és també l'equiprobable.

(b) Els valors màxims a cada columna de Π sempre valen $1/2$. Si diem $\{p_1, p_2, p_3\}$ a la distribució inicial, obtenim:

$$\overline{P}_e = 1 - p_1\frac{1}{2} + p_2\frac{1}{2} + p_3\frac{1}{2} = 1 - \frac{1}{2}(p_1 + p_2 + p_3) = 1 - \frac{1}{2} = 0.5.$$

(c) Multiplicant cada fila per la probabilitat de l'entrada corresponent, obtenim:

$$\Pi = \begin{pmatrix} 1/4 & 1/4 & 0 \\ 0 & 1/6 & 1/6 \\ 1/12 & 0 & 1/12 \end{pmatrix}.$$

Triant els valors màxims a cada columna obtenim la funció de descodificació: $B_1 \longrightarrow A_1, B_2 \longrightarrow A_1, B_3 \longrightarrow B_2$; amb probabilitat mitjana d'error:

$$\overline{P}_e = 1 - \frac{1}{4} - \frac{1}{4} - \frac{1}{6} = \frac{1}{2} - \frac{1}{6} = \frac{1}{3}.$$