# TCI - Exercicis Informació

**Exercici 1.1.** Llencem una moneda no trucada i si surt creu, acabem; si surt cara, la llencem una altra vegada i acabem. Trobeu l'entropia de la variable aleatòria associada al resultat de l'última tirada

**Exercici 1.2.** En dues ciutats, A i B, el temps pot tenir quatre estats: pluja, sol, núvols i boira. A la ciutat A les probabilitats són 1/4 per a tots els estats, i a la ciutat B són 1/4 per al sol, 1/8 per a la pluja, 1/8 per a núvol i 1/2 per a la boira. A quina ciutat fa falta més informació per a conèixer el temps?

**Exercici 1.3.** Tenim una imatge de 16 pixels x 16 pixels la qual ha estat quantificada amb 4 nivells de gris. Els nivells inicials estaven entre 0 i 255. A la imatge quantificada el nombre de pixels amb cada un dels nivells quantificats de gris és:

$$\begin{array}{rcl} \text{nivell}\,32 & \to & 32\,\text{píxels} \\ \text{nivell}\,96 & \to & 128\,\text{píxels} \\ \text{nivell}\,160 & \to & 64\,\text{píxels} \\ \text{nivell}\,224 & \to & 32\,\text{píxels} \end{array}$$

Calcula la informació mitjana (en bits/pixel) que porta la imatge quantificada

**Exercici 1. 4.** Llencem 10.000 vegades un dau perfecte i construïm un fitxer amb els 10.000 resultats obtinguts. Quanta informació és d'esperar que porti un tal fitxer?

**Exercici 1.5.** Demostreu que tot canvi de dues probabilitats en una distribució finita $p_1, \dots, p_n$ que tendeixi a apropar la distribució a una d'equidistribuïda, fa augmentar l'entropia. En concret, si $p_1 > p_2$, $0 \leq d < (p_1 - p_2)/2$ i definm $p_1' := p_1 - d$ i $p_2' := p_2 + d$ aleshores $H(p_1, p_2, \dots, p_n) \leq H(p_1', p_2', p_3, \dots, p_n)$. Indicació: apliqueu el lema de Gibbs a $q_1 := p_1'$, $q_2 := p_2'$, $q_i := p_i \,\forall i \geq 3$

**Exercici 1.6.** Volem codificar la sortida d'una FDSM amb alfabet $A := \{a, b, c, d, e\}$ i probabilitats de símbol idèntiques. Per incrementar l'eficiència codificarem per J-blocs en lloc de fer-ho per lletres individuals. Calculeu la J mínima que garanteix una eficiència superior al 99%. Doneu una estadística de font tal que la codificació per J-blocs mai tingui una eficiència superior al 30%

**Exercici 1. 7.** Considerem una FDSM amb alfabet $A := \{a_1, a_2, \dots, a_9\}$. i probabilitats .25, .25, .125, .125, .0625, .0625, .0625, .03125 i .03125. Podem codificar amb longitud fixa amb una eficiència del 100%? En cas afirmatiu, quina ha de ser la longitud dels blocs a codificar?

**Exercici 1.8.** Un microprocessador té un bus d'instruccions de 4 bits. El microprocessador admet un conjunt de 70 instruccions, 14 de les quals han de ser accedides en un sol cicle de bus. Si es vol optimitzar el temps mig d'accés a les instruccions, quantes han de ser accessibles en dos cicles de bus? Quantes han de ser accessibles en tres cicles de bus? Quantes instruccions es poden afegir al microprocessador sense variar la seva arquitectura? Nota.- El microprocessador ha de poder descodificar les instruccions instantàniament i sense ambigüitat

**Exercici 1.9.** Considerem una FDSM ternària amb alfabet $A := \{a, b, c\}$ i probabilitats respectives $0.8, 0.1, 0.1$

(a) Doneu una codificació binària que tingui una eficiència superior al 90%

(b) Si haguéssim optat per codificar amb longitud fixa, ¿de quina mida s'haurien d'haver agafat els blocs per aconseguir la mateixa eficiència?

**Exercici 1.10.** Volem codificar 40 instruccions amb un codi prefix binari. Si, per raons de disseny, hi ha d'haver exactament 3 instruccions codificades amb longitud 3 i exactament 4 instruccions codificades amb longitud 4, ¿amb quina longitud s'han de codificar les altres instruccions si volem que la longitud mitjana sigui mínima?

**Exercici 1.11.** Volem construir un codi binari prefix que contingui les paraules 0, 10 i 110. Quantes paraules addicionals de longitud 5 es poden afegir a aquest codi?

**Exercici 1. 12.** (Codis no binaris) Quants símbols calen per a construir un codi prefix amb 4 paraules de longitud 1 i 12 paraules de longitud 2?

**Exercici 1.13.** Determineu si existeix un codi binari prefix i, en aquest cas, construïu-ne un, amb paraules de longituds:

(a) 1, 2, 2, 3, 3

(b) 1, 3, 3, 3, 4, 4

(c) 2, 2, 3, 3, 4, 4, 5, 5

**Exercici 1. 14.** Doneu una codificació binària per a l'alfabet $A := \{a_1, a_2, \ldots, a_9\}$ (una paraula-codi per a cada lletra) de manera que hi hagin dues paraules-codi de longitud 2 i tres paraules-codi de longitud 3. Doneu, si és possible, una estadística de font que faci que la codificació proposada tingui una longitud mitjana igual a l'entropia. Doneu una altra estadística de font tal que la longitud mitjana de la codificació no coincideixi amb l'entropia

**Exercici 1. 15.** Construïu un codi de Huffman sobre l'alfabet A = {0, 1} a partir dels símbols i probabilitats següents:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x_1$ | 0.2 | $x_2$ | 0.18 | $x_3$ | 0.1 | $x_4$ | 0.1 |
| $x_5$ | 0.1 | $x_6$ | 0.061 | $x_7$ | 0.059 | $x_8$ | 0.04 |
| $x_9$ | 0.04 | $x_{10}$ | 0.04 | $x_{11}$ | 0.04 | $x_{12}$ | 0.03 |
| $x_{13}$ | 0.01 | | | | | | |

Compareu la longitud mitjana del codi i l'entropia de la distribució

**Exercici 1.16.** Construïu un codi de Huffman sobre l'alfabet $A := \{0, 1, 2\}$ a partir del símbols i probabilitats següents:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x_1$ | 0.3 | $x_2$ | 0.2 | $x_3$ | 0.15 | $x_4$ | 0.1 |
| $x_5$ | 0.1 | $x_6$ | 0.08 | $x_7$ | 0.05 | $x_8$ | 0.02 |

**Exercici 1. 17.** Una font sense memòria emet símbols $a, b, c, d$ amb probabilitats $0.4, 0.3, 0.2, 0.1$ respectivament. Trobeu una codificació binària que no requereixi més de 1.87 bits per símbol

**Exercici 1.18.** Suposem una FDSM amb alfabet $a := \{a, b, c, d, e\}$ i probabilitats respectives 0.5, 0.2, 0.1, 0.1 i 0.1. Codifiqueu per Huffman lletra a lletra i per blocs de longitud 2 i compareu els resultats

# TCI - Solucions Informació

**Solució 1.1.** 0.811 bits

**Solució 1.2.** A la ciutat A

**Solució 1.3.** 1.75 bits/pixel

**Solució 1.4.** 4308.27 bits

**Solució 1.5.** Pel lema de Gibbs, $-\sum_{i\geq 1} p_i \cdot \log_2 p_i \leq -p_1 \cdot \log_2(p'_1) - p_2 \cdot \log_2(p'_2) - \sum_{i\geq 3} p_i \cdot \log_2 p_i$, és a dir

$$-p_1 \cdot \log_2(p_1) - p_2 \cdot \log_2(p_2) \leq -p_1 \cdot \log_2(p'_1) - p_2 \cdot \log_2(p'_2)$$

Però $d \cdot \log_2 \frac{p'_1}{p'_2} \geq 0$. Per tant,

$$-p_1 \cdot \log_2(p_1) - p_2 \cdot \log_2(p_2) \leq -p_1 \cdot \log_2(p'_1) - p_2 \cdot \log_2(p'_2) + d \cdot \log_2 p'_1 - d \cdot \log_2 p'_2$$

és a dir,

$$-p_1 \cdot \log_2(p_1) - p_2 \cdot \log_2(p_2) \leq -p'_1 \cdot \log_2(p'_1) - p'_2 \cdot \log_2(p'_2)$$

Per tant, $H(p_1, p_2, p_3, \ldots, p_n) \leq H(p'_1, p'_2, p_3, \ldots, p_n)$

**Solució 1.6.** $J_{\min} = 29$; per exemple: $p_a := 0.9$, $p_b = p_c = p_d = p_e := 0.025$

**Solució 1.7.** No és possible codificar amb longitud fixa al 100% d'eficiència

**Solució 1.8.** 30 instruccions han de ser accessibles en dos cicles de bus i 26 en tres cicles de bus. Al microprocessador se li podrien afegir 6 instruccions més (accessibles en tres cicles de bus)

**Solució 1.9.** Per exemple:

$$\begin{array}{lll} aa \to 0 & ba \to 1010 & ca \to 1110 \\ ab \to 100 & bb \to 10110 & cb \to 11110 \\ ac \to 110 & bc \to 10111 & cc \to 11111 \end{array}$$

Amb longitud fixa (copdificant per blocs) no es pot obtebuir una eficiència superior al 90%

**Solució 1.10.** S'han de codificar 0 instruccions amb longitud 5, 15 instruccions amb longitd 6 i 18 instruccions amb longitud 7

**Solució 1.11.** 4

**Solució 1.12.** 6

**Solució 1. 13.** 1: NO existeix: 2: (per exemple) $0, 100, 101, 110, 1110, 1111$. 3: (per exemple) $00, 01, 100, 101, 1100, 1101, 11100, 11110$

**Solució 1.14.** Codificació:

$$\begin{array}{lll} a_1 \to 00 & a_4 \to 101 & a_7 \to 11101 \\ a_2 \to 01 & a_5 \to 110 & a_8 \to 11110 \\ a_3 \to 100 & a_6 \to 11100 & a_9 \to 11111 \end{array}$$

Si $p_1 = p_2 := 1/4$, $p_3 = p_4 = p_5 := 1/8\, p_6 = p_7 = p_8 = p_9 := 1/32$ llavors la longitud mitjana del codi coincideix amb l'entropia. I si, per exemple $p_1 = \ldots = p_9 = 1/9$ llavors lo longitud mitjana del codi no coincideix amb l'entropia

**Solució 1.15.** Per exemple:

$$\begin{array}{llllllll} x_1 \to & 00 & x_2 \to & 110 & x_3 \to & 010 & x_4 \to & 011 \\ x_5 \to & 1110 & x_6 \to & 1000 & x_7 \to & 1001 & x_8 \to & 1010 \\ x_9 \to & 10110 & x_{10} \to & 10111 & x_{11} \to & 11110 & x_{12} \to & 111110 \\ x_{13} \to & 111111 & & & & & & \end{array}$$

$\bar{\ell} = 3.42$ $H = 3.35$

**Exercici 1.16.** Per exemple:

$$\begin{array}{llllllll} x_1 \to & 00 & x_2 \to & 10 & x_3 \to & 011 & x_4 \to & 110 \\ x_5 \to & 111 & x_6 \to & 0110 & x_7 \to & 01110 & x_8 \to & 01111 \end{array}$$

**Solució 1.17.** Codificació de Huffman amb blocs de longitud 2. Per exemple:

$$\begin{array}{llllllll} aa \to & 000 & ba \to & 010 & ca \to & 1010 & da \to & 00111 \\ ab \to & 100 & bb \to & 111 & cb \to & 1011 & db \to & 11011 \\ ac \to & 0010 & bc \to & 0110 & cc \to & 00110 & dc \to & 011110 \\ ad \to & 1110 & bd \to & 01110 & cd \to & 11010 & dd \to & 011111 \end{array}$$

**Solució 1.18.** Les codificacions les feu vosaltres. L'eficiència passa del 98.05% al 98.79%