Teoría de la información

breve introducción práctica

Javier García, Graphics & Imaging Laboratory (UdG)

Contenido

- Introducción
- Definición de Información y Entropía
- Ejemplos de aplicación I
- Ejemplos de aplicación II

Introducción

- Entropía física (siglo XIX) vs Entropía de información (siglo XX)
- Telecomunicaciones & informática basada en T.Información

Información & Entropía

Ω = conjunto de posibles resultados de una observación

p_i = probabilidad de un resultado concreto

Valor promedio de $x = \sum_i p_i x_i$

I_i = información asociada al resultado i $\longrightarrow$ $I_i = -\log_2 p_i$

Ejemplo 1: Lanzar una moneda

$$\Omega = \{cara, cruz\}$$

$$p_{cara} = \tfrac{1}{2} \quad p_{cruz} = \tfrac{1}{2}$$

$$I_{cara} = -\log_2 \tfrac{1}{2} = 1 \text{ bit}$$

$$I_{cruz} = -\log_2 \tfrac{1}{2} = 1 \text{ bit}$$

$$\boxed{-\sum_i p_i \log_2 p_i = -\tfrac{1}{2} \log_2 \tfrac{1}{2} - \tfrac{1}{2} \log_2 \tfrac{1}{2} = 1 \text{ bit}}$$

Información & Entropía

Ejemplo 2 : Género de personas viendo el fútbol en un bar

$$\Omega = \{\text{hombre}, \text{mujer}\}$$

$$p_{\text{hombre}} = \frac{9}{10} \quad p_{\text{mujer}} = \frac{1}{10}$$

$$I_{\text{hombre}} = -\log_2 \frac{9}{10} = 0.15 \text{ bits}$$

$$I_{\text{mujer}} = -\log_2 \frac{1}{10} = 3.32 \text{ bits}$$

$$\boxed{-\sum_i p_i \log_2 p_i = -\frac{9}{10} \log_2 \frac{9}{10} - \frac{1}{10} \log_2 \frac{1}{10} = 0.47 \text{ bit}}$$

Ejemplo 3 : Moneda con caras iguales

$$\Omega = \{\text{cara}, \text{cruz}\}$$

$$p_{\text{cara}} = 1 \quad p_{\text{cruz}} = 0$$

$$I_{\text{cara}} = -\log_2 1 = 0 \text{ bits}$$

$$I_{\text{cruz}} = -\log_2 0 = \infty \text{ bits}$$

$$\boxed{-\sum_i p_i \log_2 p_i = -1 \log_2 1 - 0 \log_2 0 = 0 \text{ bit}}$$

Información & Entropía

Ejemplo 1: Lanzar una moneda 1 bit

Ejemplo 2 : Género de personas viendo el fútbol en un bar 0.47 bits

Ejemplo 3 : Moneda con caras iguales 0 bits

Todos los experimentos expuestos en los ejemplos anteriores tienen **2** posibles resultados.

¿Qué aprendemos?

- Cuando los resultados son equiprobables: **1 bit de información en promedio**
- Cuando sólo un resultado es posible: **0 bits de información en promedio**
- Cualquier otra posibilidad: **entre 0 y 1 bits de información**.

ENTROPIA ~ INFORMACION PROMEDIO EN EL RESULTADO

Ejemplos de aplicación I

Problema a solucionar

Una máquina ‘expulsa’ números X.

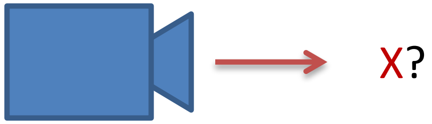
Posibles salidas de X: **1, 2, 3, 4, 5, 6, 7 ó 8**

Probabilidades:

$p_1, p_2, p_3, p_4, p_5, p_6, p_7$ y p_8

Queremos saber qué número ha salido.

Solo podemos hacer preguntas de **SÍ** o **NO**



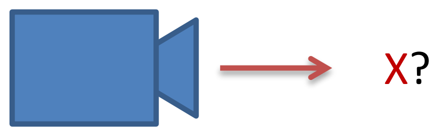
¿Cuál es el mínimo número de preguntas?

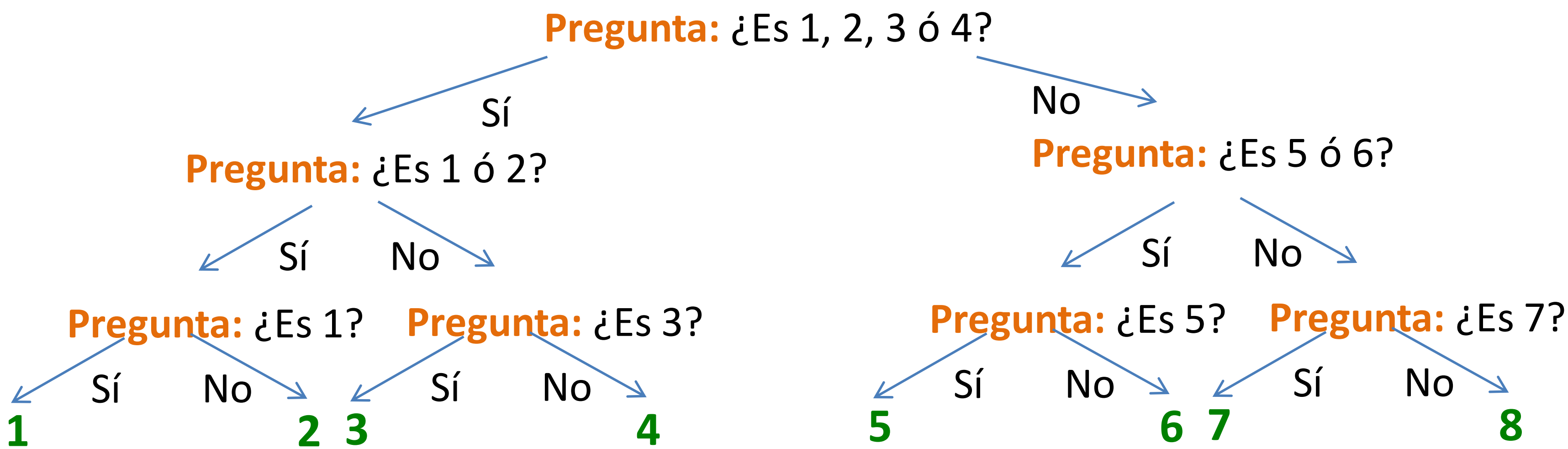
Ejemplos de aplicación I

(http://blog.pseudolog.com)

CASO 1: Búsqueda binaria

$$p_1 = p_2 = \ldots = p_8 = \frac{1}{8}$$





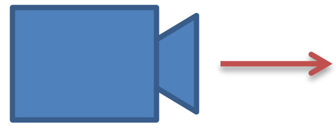
NUMERO DE PREGUNTAS: 3

Ejemplos de aplicación I

CASO 2: Búsqueda Jerárquica

$p_1 = 0.35 \quad p_2 = 0.25 \quad p_3 = 0.12 \quad p_4 = 0.10$

$p_5 = 0.08 \quad p_6 = 0.06 \quad p_7 = 0.03 \quad p_8 = 0.01$



X?

Pregunta: ¿Es 1?
Sí No
1
Pregunta: ¿Es 2?
Sí No
2
Pregunta: ¿Es 3?
Sí No
3
Pregunta: ¿Es 4?
Sí No
4
Pregunta: ¿Es 5?
Sí No
5
Pregunta: ¿Es 6?
Sí No
6
Pregunta: ¿Es 7?
Sí No
7 **8**

Valor esperado de número de preguntas

$$p_1 \times 1 + p_2 \times 2 + \ldots + p_7 \times 7$$

$$0.35 \times 1 + 0.25 \times 2 + 0.12 \times 3 + 0.10 \times 4 + 0.08 \times 5 + 0.06 \times 6 + 0.03 \times 7 = 2.83$$

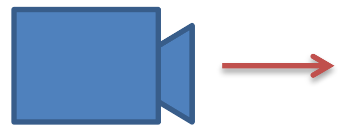
NUMERO DE PREGUNTAS: 2.83

Ejemplos de aplicación I

CASO 3: SHANNON-FANO

ELEGIR PREGUNTAS ~ BINARIAS en probabilidades

$$p_1 = 0.35 \quad p_2 = 0.25 \quad p_3 = 0.12 \quad p_4 = 0.10$$
$$p_5 = 0.08 \quad p_6 = 0.06 \quad p_7 = 0.03 \quad p_8 = 0.01$$



X?

Pregunta: ¿Es 1,3 ó 7?

$$p_1 + p_3 + p_7 = 0.5 \mid p_2 + p_4 + p_5 + p_6 + p_8 = 0.5$$

Sí

Pregunta: ¿Es 1?

$$p_1 = 0.35 \mid p_3 + p_7 = 0.15$$

Sí **1**

No

Pregunta: ¿Es 3?

$$p_3 = 0.12 \mid p_7 = 0.03$$

Sí **3**

No **7**

No

Pregunta: ¿Es 2?

$$p_2 = 0.25 \mid p_4 + p_5 + p_6 + p_8 = 0.25$$

Sí **2**

No

Pregunta: ¿Es 5 ó 6?

$$p_5 + p_6 = 0.14 \mid p_4 + p_8 = 0.11$$

Sí

Pregunta: ¿Es 5?

$$p_5 = 0.08 \mid p_6 = 0.06$$

Sí **5**

No **6**

No

Pregunta: ¿Es 4?

$$p_4 = 0.10 \mid p_8 = 0.01$$

Sí **4**

No **8**

1 preguntas. Probabilidad = 0
2 preguntas. Probabilidad = 0.35 + 0.25 = 0.60
3 preguntas. Probabilidad = 0.12 + 0.03 = 0.15
4 preguntas. Probabilidad = 0.08 + 0.06 + 0.10 + 0.01 = 0.25

Valor esperado de preguntas = 1 × 0 + 2 × 0.60 + 3 × 0.15 + 4 × 0.25 = 2.65

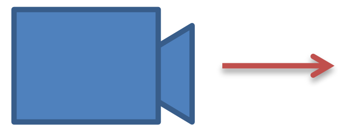
NUMERO DE PREGUNTAS: 2.65

Ejemplos de aplicación I

CASO 4: HUFFMAN (óptimo)

CONSTRUIR ARBOL DE MENOR A MAYOR PROBABILIDAD

$$p_1 = 0.35 \quad p_2 = 0.25 \quad p_3 = 0.12 \quad p_4 = 0.10$$
$$p_5 = 0.08 \quad p_6 = 0.06 \quad p_7 = 0.03 \quad p_8 = 0.01$$



X?

Pregunta: ¿Es 1 ó 2? **1**

Sí

Pregunta: ¿Es 1? **0.60**

Sí No

1 0.35 **2 0.25**

No

Pregunta: ¿Es 3 ó 4? **0.40**

Sí

Pregunta: ¿Es 3? **0.22**

Sí No

3 0.12 **4 0.10**

No

Pregunta: ¿Es 5? **0.18**

Sí No

5 0.08

Pregunta: ¿Es 6? **0.10**

Sí No

6 0.06

Pregunta: ¿Es 7? **0.04**

Sí No

7 0.03 **8 0.01**

$$(0.35 + 0.25) * 2 + (0.12 + 0.10 + 0.08) * 3 + 0.06 * 4 + (0.03 + 0.01) * 5 = 2.54$$

NUMERO DE PREGUNTAS: 2.54

Ejemplos de aplicación I

HUFFMAN (óptimo) versus ENTROPIA

Codificación de las preguntas en **0**'s y **1**'s

- **Pregunta:** ¿Es 1 ó 2?
  - Sí **1** — **Pregunta:** ¿Es 1?
    - Sí **1** → **1**
    - No **0** → **2**
  - No **0** — **Pregunta:** ¿Es 3 ó 4?
    - Sí **1** — **Pregunta:** ¿Es 3?
      - Sí **1** → **3**
      - No **0** → **4**
    - No **0** — **Pregunta:** ¿Es 5?
      - Sí **1** → **5**
      - No **0** — **Pregunta:** ¿Es 6?
        - Sí **1** → **6**
        - No **0** — **Pregunta:** ¿Es 7?
          - Sí **1** → **7**
          - No **0** → **8**

Codificación	probabilidad	información
1 = 11	$p_1 = 0.35$	$-\log_2 p_1 = 1.8646$
2 = 10	$p_2 = 0.25$	$-\log_2 p_2 = 2.25$
3 = 011	$p_3 = 0.12$	$-\log_2 p_3 = 3.1789$
4 = 010	$p_4 = 0.10$	$-\log_2 p_4 = 3.4219$
5 = 001	$p_5 = 0.08$	$-\log_2 p_5 = 3.7239$
6 = 0001	$p_6 = 0.06$	$-\log_2 p_6 = 4.1189$
7 = 00001	$p_7 = 0.03$	$-\log_2 p_7 = 5.0889$
8 = 00000	$p_8 = 0.01$	$-\log_2 p_8 = 6.6539$

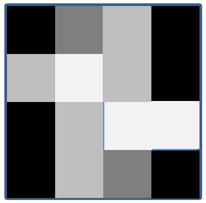
Cota inferior de número de preguntas

$$\text{ENTROPIA} = -(p_1\log_2 p_1 + p_2\log_2 p_2 + p_3\log_2 p_3 + p_4\log_2 p_4 + p_5\log_2 p_5 + p_6\log_2 p_6 + p_7\log_2 p_7 + p_8\log_2 p_8) = 2.4826$$

Ejemplos de aplicación II

COMPRESION IMAGENES SIN PERDIDAS CON HUFFMAN

IMAGEN DE 4x4 píxeles
con 2 bits/pixel -> 4 niveles de gris



Nivel de gris	FRECUENCIAS DE APARICIÓN	PROBABILIDADES DE APARICIÓN
00	6	$\frac{6}{16} = 0.375$
01	2	$\frac{2}{16} = 0.125$
10	5	$\frac{5}{16} = 0.3125$
11	3	$\frac{3}{16} = 0.1875$

00	0.375	0.375	0.375	0.625
10	0.3125	0.3125	0.625	0.375
11	0.1875	0.3125		
01	0.125			

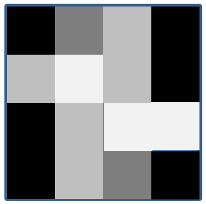
00	0.375 1	0.375 1	0.625	0
10	0.3125 00	0.3125 00	0.375	1
11	0.1875 010	0.3125 01		
01	0.125 011			

Original		Huffman
00	→	1
10	→	00
11	→	010
01	→	011

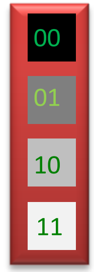
Ejemplos de aplicación II

COMPRESION IMAGENES SIN PERDIDAS CON HUFFMAN

IMAGEN DE 4x4 píxeles
con 2 bits/pixel -> 4 niveles de gris



Nivel de gris	FRECUENCIAS DE APARICIÓN	PROBABILIDADES DE APARICIÓN
00	6	$\frac{6}{16} = 0.375$
01	2	$\frac{2}{16} = 0.125$
10	5	$\frac{5}{16} = 0.3125$
11	3	$\frac{3}{16} = 0.1875$



Original		Huffman
00	$\longrightarrow$	1
10	$\longrightarrow$	00
11	$\longrightarrow$	010
01	$\longrightarrow$	011

Original

0001100010111000001011100100100

Huffman

10110010001000110001001010001111

Original $L = 2 \times 0.375 + 2 \times 0.3125 + 2 \times 0.1875 + 2 \times 0.125 = 2$ bits/pixel

Huffman $L = 1 \times 0.375 + 2 \times 0.3125 + 3 \times 0.1875 + 3 \times 0.125 = 1.9375$ bits/pixel

ENTROPIA

$$-(0.375 \times \log_2 0.375 + 0.3125 \times \log_2 0.3125 + 0.1875 \times \log_2 0.1875 + 0.125 \times \log_2 0.125) = 1.8829 \text{ bits/pixel}$$

Telecomunicaciones
Basadas en teoría de la información

Gracias por la atención