Teoría de la Información

→ cota inferior al nro de bits necesarios para codificar un mensaje.

→ lo usaremos para cotas inferiores de algoritmos basados en comparaciones

→ Las respuestas a las comparaciones que hace un algoritmo sirven como una codificación de su input/output, y x teoría de la información no pueden ser demasiado pocas.

→ permite cotas inferiores de peor caso y de caso promedio.

Peor caso

para codificar un elemento de un conjunto U se necesitan al menos

$\boxed{\log_2 |U| \text{ bits}}$ en el peor caso

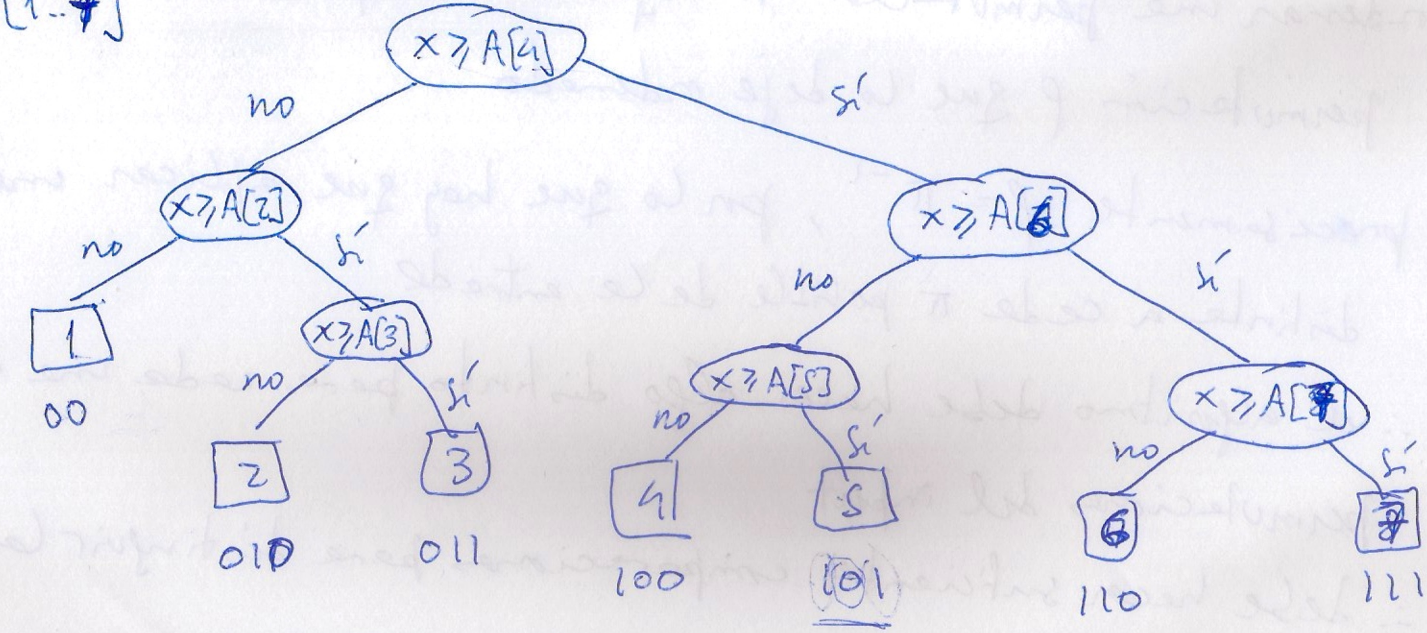
→ el total de códigos distintos con menos de $\log_2 |U|$ bits es

$$\sum_{l=0}^{\lceil \log_2 |U| \rceil - 1} 2^l = 2^{\lceil \log_2 |U| \rceil} - 1 < |U|$$

si $|U|$ potencia de 2.

Ej ~~búsqueda lineal~~ búsqueda en un arreglo ordenado (por comparaciones)

A[1..7]



× T.I. en el peor caso tengo que hacer a lo menos

$\log_2 7$ ~~propor~~ comparaciones → 3 comparaciones

para A[1..n], en el peor caso necesito $\log_2 n$ comparaciones.

Ordenar por comparaciones $A[1..n]$

- ordenar una permutación π equivale a aplicarle al input una permutación ρ que lo deje ordenado
- precisamente $\rho = \pi^{-1}$, por lo que hay que aplicar una permutación distinta a cada π posible de la entrada
- $\therefore$ el algoritmo debe hacer algo distinto para cada una de las $n!$ permutaciones del input
- $\therefore$ debe hacer suficientes comparaciones para distinguirlas.

× T. I. debe hacer al menos $\log_2 n!$ comparaciones.

(Stirling) $n! = \frac{n^n}{e^n}\sqrt{2\pi n}\,(1+O(1/n))$

$$\log_2 n! = n\log_2 n - \log_2 e + O(\log n)$$
$$= n\log_2 n - \Theta(n) = \Omega(n\log n)$$

cota inferior para ordenar n elementos por comparaciones

Caso promedio

(Kolmogorov)

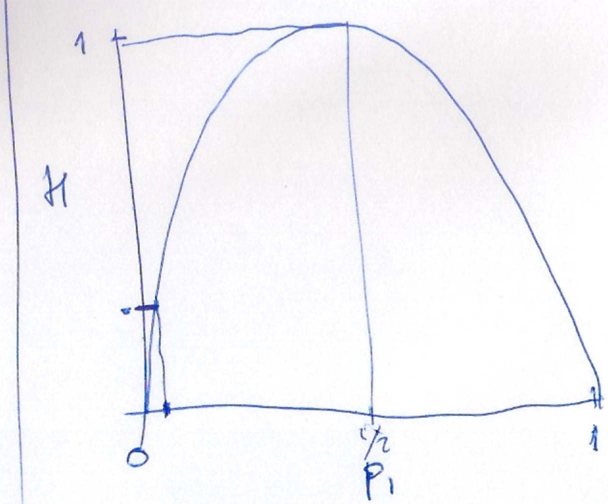
Teorema de Shannon (1948)

Para distinguir un elemento de un conjunto de n elementos que aparecen con probabilidades $p_1, \dots, p_n$, se necesitan en promedio

$$\log_2 n \leq H = \sum p_i \boxed{\log_2 \frac{1}{p_i}} \text{ bits.}$$

ej p_1, p_2 $\quad p_1 + p_2 = 1$

$$H = p_1 \log \frac{1}{p_1} + p_2 \log \frac{1}{p_2}$$



Obs: si todos los p_i son $p_i = 1/n$

entonces $H = \sum \frac{1}{n} \log_2 n = \log_2 n$ bits.

→ las cotas inferiores que obtenemos por teoría de la información en el peor caso también valen en promedio si todos los inputs son igualmente probables.

→ buscar en un arreglo ordenado cuesta $\Omega(\log n)$ en promedio

→ ordenar cuesta $\Omega(n \log n)$ en promedio.