Teoría de Información y Codificación (66.24 — 86.11)

Práctica 3

1. Dado el código binario $(a, b, c) \to (0, 10, 011)$ ¿es no singular? ¿es unicamente decodificable? ¿es instantáneo?

2. Dada una fuente de n símbolos con probabilidades $p_1 \cdots p_n$

   a) Explicar por qué existe un código instantáneo con longitudes $l_i = \lceil -\log p_i \rceil$.

   b) Demosrtar que la longitud media de este código cumple $H \le L < H + 1$.

   c) ¿Implica lo anterior que las longitudes de este código ℓ_i difieren de las longitudes del código realizable óptimo ℓ_i^o en no más de uno? ($l_i < \ell_i^o + 1$?)

   d) Considerar el caso particular $n = 2$, $p = (0.1, 0.9)$ (calcular entropía, encontrar código del punto a y el código óptimo)

3. Una fuente emite con tres símbolos de probabilidades $\mathbf{p} = (0.2, 0.3, 0.5)$. Encontrar el código binario óptimo y calcular la longitud promedio. Idem para extensión de orden 2 (suponiendo que los símbolos son independientes) . ¿Puede afirmarse algo a priori sobre la longitud promedio de extensiones superiores?

4. Sea una fuente de cinco simbolos con las probabilidades $(0.38, 0.18, 0.15, 0.15, 0.14)$

   a) Hallar la entropía en base 2 (bits) y en base 3.

   b) Calcular el código de Huffman binario y el código de Shannon-Fano. ¿Qué se puede afirmar a priori sobre las longitudes promedio L_H L_{SF}? Calcular los valores.

   c) Calcular el códiigo de Huffman ternario y la longitud promedio de este código.

5. Una fuente emite 5 símbolos según esta ley: repite el símbolo anterior con probabilidad 1/2, sino emite uno de los restantes equiprobablemente. El primer símbolo se elije con probabilidad uniforme. Calcular la tasa de entropía y proponer un código.

6. *a*) ¿Es posible armar un código instantáneo binario con longitudes $(1, 2, 3, 3, 4)$?

   b) Un código binario tiene longitudes $(1, 2, 3, 4)$ ¿es instantáneo?

7. Una fuente de cuatro símbolos estacionaria tiene entropia de $H(X)$ bits para cada símbolo, la tasa de entropía es H_r En cada uno de los casos siguientes, acotar (si es posible, calcular) L, la longitud media del código de Huffman sin usar extensiones de la fuente (codificamos cada simbolo por separado). Analizar si es esperable una reducción en L por simbolo, si se toman extensiones de orden superior. ¿Qué pasaría si se usa codificación aritmética?

   a) $H(X) = 2$, $H_r = 2$

   b) $H(X) = 2$, $H_r = 1$

   c) $H(X) = 1.8$, $H_r = 1.8$

8. Simular las fuentes de ej.3, 4, 5,7 y estudiar qué sucede al comprimir con un compresor cualquiera.

9. Mostrar que la entropía relativa mide el costo de codificar con una distribución de probabilidad falsa: Ejercicio Cover y Thomas 2ed: 5.30.

10. Mostrar que si un código cumple la desigualdad de Kraft estricta, entonces no todas las secuencias posibles (entradas al decodificador) representan una codificación válida. (ref: CT ej 5.3)

11. Analizar esta afirmación. Un código binario instantáneo $Y = C(X)$ alcanza la cota inferior de Shannon si y solo si la secuencia de salida tiene tasa de entropía $H_r(Y) = 1$

12. Ejercicios CT: 5.14, 5.15, 5.16, 5.20, 5.21, 5.38

13. Ejercicio CT: 5.35 Generar una variable aleatoria Bernoulli con p arbitraria usando una moneda equilibrada.

14. Codificar con LZ78 la secuencia AAAAAABBABABAAAAABBABAB