

Tema 4: Teoría de la Información.

Problemas de Fuentes de Información

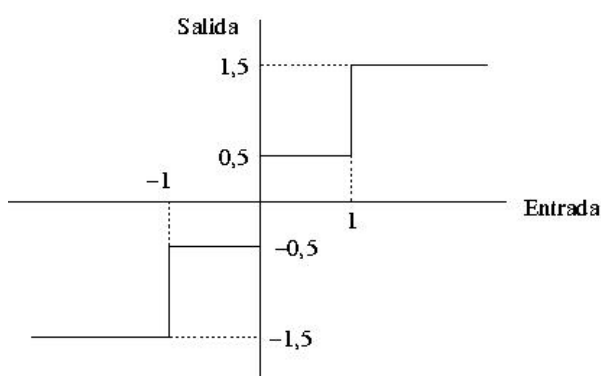
1. Una fuente de mensajes genera uno de los cuatro mensajes de forma aleatoria cada microsegundo. Las probabilidades de estos mensajes son 0.4, 0.3, 0.2 y 0.1. Cada mensaje emitido es independiente de cualquier otro mensaje en la secuencia.
 - a. ¿Cuál es la entropía de la fuente?
 - b. ¿Cuál es la tasa de información generada por esta fuente (en bits por segundo)
 - c. Comparar los resultados obtenidos con los que se obtendrían si los mensajes fueran equiprobables.
2. Una imagen de TV blanco y negro de alta resolución constan de $2 \cdot 10^6$ píxeles y 16 niveles de luminancia. Las imágenes se repiten a un razón de 32 por segundo. Se supone que todos los píxeles son independientes entre sí y todos los niveles de luminancia tienen la misma probabilidad de ocurrencia. Calcular la tasa de información promedio emitida por esta fuente de imágenes de TV.
3. Considere la una fuente como la telegráfica que tiene dos símbolos, punto y raya. La duración de un punto es 0.2 segundos y la de la raya es tres veces la del punto. La probabilidad de que ocurra un punto es dos veces la de la raya y el tiempo entre símbolos es de 0.2 segundos. Calcular la tasa de información de la fuente telegráfica.
4. Dos variables aleatorias binarias X e Y están distribuidas de acuerdo a la siguiente distribución conjunta: $P(X=Y=0)=P(X=0, Y=1)= P(X=Y=1)=1/3$. Calcular $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$ Y $H(X,Y)$.

5. Una fuente sin memoria tiene el alfabeto $\mathcal{A} = \{-5, -3, -1, 0, 1, 3, 5\}$ con las probabilidades correspondientes $\{0.05, 0.1, 0.1, 0.15, 0.05, 0.25, 0.3\}$.
- ¿Cuál es la entropía de la fuente?
 - Suponga que la fuente se cuantiza de acuerdo a la siguiente regla de cuantización:

$$\begin{aligned} \begin{cases} -5 \\ -3 \\ -1 \\ 0 \\ 1 \\ 3 \\ 5 \end{cases} &= \begin{cases} -5 \\ -3 \\ -1 \\ 0 \\ 1 \\ 3 \\ 5 \end{cases} \\ &= \begin{cases} -5 \\ -3 \\ -1 \\ 0 \\ 1 \\ 3 \\ 5 \end{cases} \\ &= \begin{cases} -5 \\ -3 \\ -1 \\ 0 \\ 1 \\ 3 \\ 5 \end{cases} \end{aligned}$$

¿Cuál es la entropía de la fuente cuantizada?

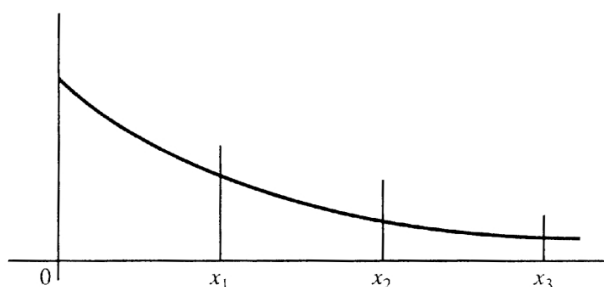
6. Un proceso gaussiano de media cero y varianza unidad se muestrea y las muestras se aplican a un cuantizador uniforme que tiene la característica de la figura. Calcular la entropía de la salida del cuantizador.



7. La entrada a un cuantizador es una variable aleatoria que tiene una función densidad de probabilidad dada por:

$$f(x) = \begin{cases} 3 - x & 0 \leq x < 3 \\ 0 & \text{en otro caso} \end{cases}$$

La señal se va a cuantizar con un cuantizador de 4 niveles como se muestra en la figura. Determinar los valores de x_i , $i=1, 2, 3$ en términos de 'a' de forma que la entropía a la salida del cuantizador sea máxima.



- 8.** Una señal se muestrea 500 veces por segundo y las muestras tienen una función densidad de amplitudes gaussiana de media cero y varianza S^2 . Las muestras se cuantizan de acuerdo a la tabla siguiente.
- Determinar la entropía de la salida del cuantizador y la tasa de información en bits por segundo.
 - Determinar los intervalos de cuantización en función de S de forma que a la salida del cuantizador la entropía sea máxima.

Entrada al cuantizador	Salida del cuantizador
$-\infty < x_i < -2S$	m_0
$-2S < x_i < -S$	m_1
$-S < x_i < 0$	m_2
$0 < x_i < S$	m_3
$S < x_i < 2S$	m_4
$2S < x_i < \infty$	m_5

- 9.** Una fuente DMS tiene 5 símbolos x_1, x_2, x_3, x_4 y x_5 con probabilidades $P(x_1)=0.4$, $P(x_2)=0.19$, $P(x_3)=0.16$, $P(x_4)=0.15$, $P(x_5)=0.1$. Construir un código de Huffman y calcular la eficiencia del código.

- 10.** Una fuente emite siete mensajes con probabilidades $1/2, 1/4, 1/8, 1/16, 1/32, 1/64$ y $1/64$, respectivamente.

- Encuentre la entropía de la fuente. Obtenga un código compacto binario y calcule la longitud promedio de las palabras código. Determinar la eficiencia del código.

- 11.** Una fuente emite siete mensajes con probabilidades $1/3, 1/3, 1/9, 1/9, 1/27, 1/27$ y $1/27$, respectivamente.

- Encuentre la entropía de la fuente. Obtener un código compacto binario, calcular la longitud promedio de las palabras código y determinar la eficiencia del código.

- 12.** Una fuente emite aleatoriamente tres mensajes equiprobables e independientes.
- a.** Encuentre la entropía de la fuente.
 - b.** Obtenga un código compacto binario y calcule la longitud promedio de las palabras código. Determinar la eficiencia del código.
 - c.** Para mejorar la eficiencia del código binario, se codifica la extensión segunda de la fuente. Obtener un código compacto binario, calcular la longitud promedio de las palabras código y determinar la eficiencia del código.
- 13.** Un alfabeto de entrada consta de 100 caracteres.
- a.** Si cada carácter se codifica con un código de longitud fija, determinar el número necesario de bits para codificarlos.
 - b.** Vamos a hacer la suposición de que 10 caracteres son igualmente probables y que cada uno ocurre con una probabilidad de 0.05. Vamos a suponer que el resto de los 90 caracteres son igualmente probables. Determinar el número de bits necesarios para codificar este alfabeto utilizando un código Huffman de longitud variable.

1. Una fuente de mensajes genera uno de los cuatro mensajes de forma aleatoria cada microsegundo. Las probabilidades de estos mensajes son 0.4, 0.3, 0.2 y 0.1. Cada mensaje emitido es independiente de cualquier otro mensaje en la secuencia.

Atun

- a. ¿Cuál es la entropía de la fuente?
- b. ¿Cuál es la tasa de información generada por esta fuente (en bits por segundo)?
- c. Comparar los resultados obtenidos con los que se obtendrían si los mensajes fueran equiprobables.

Los mensajes serán :

$$* X_1 \rightarrow P(X_1) = 0.4$$

$$* X_2 \rightarrow P(X_2) = 0.3$$

$$* X_3 \rightarrow P(X_3) = 0.2$$

$$* X_4 \rightarrow P(X_4) = 0.1$$

A) La entropía viene determinado por:

$$H(X) = \sum_{x \in X} p_x \log_2 \left(\frac{1}{p_x} \right) = - \sum_{x \in X} p_x \log_2 (p_x) =$$

$$= - (0.4 \log_2(0.4) + 0.3 \log_2(0.3) + 0.2 \log_2(0.2) + 0.1 \log_2(0.1)) =$$

$$= 1.85 \text{ bits/sym}$$

$$B) R(X) = r \cdot H(X) = 1.85 \cdot 10^{-6} \text{ bits/s}$$

↑

del enunciado: $1 \mu s = 10^{-6} s$

C) Si $X_1 \dots X_4$ tienen $P(X_1) = \dots = P(X_4) = 0.25$:

$$H'(X) = - 4 \cdot 0.25 \log_2(0.25) = 2 \text{ bits/sym}$$

$$R'(X) = 2 \cdot 10^{-6} \text{ bits/s}$$

2. Una imagen de TV blanco y negro de alta resolución constan de $2 \cdot 10^6$ píxeles y 16 niveles de luminancia. Las imágenes se repiten a un razón de 32 por segundo. Se supone que todos los píxeles son independientes entre sí y todos los niveles de luminancia tienen la misma probabilidad de ocurrencia. Calcular la tasa de información promedio emitida por esta fuente de imágenes de TV.



$2 \cdot 10^6$ píxeles con 16 posibles valores de luminancia:

→ mensajes: $l_1, \dots, l_{16}$ con $P(l_1) = \dots = P(l_{16}) = \frac{1}{16}$

→ $H(l) = -16 \cdot \frac{1}{16} \log_2(1/16) = 4 \text{ bits/sym}$

→ $r = 32 \text{ imgs/s} \cdot 2 \cdot 10^6 \text{ sym/img} = 6,4 \cdot 10^7 \text{ sym/s}$

Entonces $R(x) = r \cdot H(x) = 2,56 \cdot 10^8 \text{ bits/s}$

3. Considere la una fuente como la telegráfica que tiene dos símbolos, punto y raya
- La duración de un punto es 0.2 segundos y la de la raya es tres veces la del punto. La probabilidad de que ocurra un punto es dos veces la de la raya y el tiempo entre símbolos es de 0.2 segundos. Calcular la tasa de información de la fuente telegráfica.

Atun

2 mensajes: x_1 (punto) $P(x_1) = 2 P(x_2) = 1/3$

x_2 (raya) $P(x_2) = (2) = 2/3$



$$P(x_1) + P(x_2) = 1$$

$$2P(x_2) + P(x_2) = 1 \Leftrightarrow P(x_2) = 1/3$$

Calculo la entropía:

$$H(x) = - \frac{2}{3} \log_2 (2/3) - \frac{1}{3} \log_2 (1/3) = 0.92 \text{ bits/sym}$$

Calculo r:

Sabemos que punto dura 0.2s

raya dura 0.6s

distancia entre sym 0.2s

} Calculo el tiempo promedio para obtener cuánto dura un mensaje:

$$\bar{t} = (0.2s + 0.2s) \cdot \frac{2}{3} + (0.6 + 0.2) \cdot \frac{1}{3} = 0.53$$



Sumamos el tiempo que se tarda hasta enviar otro sym.

$$R(x) = \frac{H(x)}{\bar{t}} = 1.73 \text{ bits/s}$$

4. Dos variables aleatorias binarias X e Y están distribuidas de acuerdo a la siguiente distribución conjunta: $P(X=Y=0)=P(X=0, Y=1)=P(X=Y=1)=1/3$.
Calcular $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$ y $H(X,Y)$.

Atun

$$P(X=Y=0) = P(X=0, Y=1) = P(X=Y=1) = 1/3.$$

Calcular $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$ y $H(X,Y)$:

De aquí saco que $P(X=0, Y=0) = P(X=0, Y=1) = 1/3$, es decir:

$$P(X=0) = 2/3; \text{ y que } P(X=1, Y=1) = 1/3 \text{ y } P(X=1, Y=0) = 0.$$

Entonces:

$$H(X) = -\frac{1}{3} \log_2(1/3) - \frac{2}{3} \log_2(2/3) = 0.92 \text{ bits/sym}$$

Saco que $P(Y=0, X=0) = 1/3$ y $P(Y=0, X=1) = 0$ y que

$$P(Y=1, X=1) = P(Y=1, X=0) = 1/3, \text{ entonces:}$$

$$H(Y) = -\frac{1}{3} \log_2(1/3) - \frac{2}{3} \log_2(2/3) = 0.92 \text{ bits/sym}.$$

Con $H(X,Y)$ podré despejar $H(Y|X)$ y $H(X|Y)$:

$$H(X,Y) = \sum_i \sum_j p(x_i, y_j) \log_2 \left[\frac{1}{p(x_i, y_j)} \right] = - \sum_i \sum_j p(x_i, y_j) \log_2(p(x_i, y_j)) = 1.58 \text{ bits/sym}$$

$$H(X,Y) = -[p(X=0, Y=0) \log_2(p(X=0, Y=0)) + \dots + p(X=1, Y=1) \log_2(\dots)]$$

$$H(X,Y) = -3 p(X=0, Y=0) \log_2(p(X=0, Y=0)) =$$

→ puedo hacer esto por que ● tienen misma prob.

$$H(X,Y) = -\log_2(p(X=0, Y=0)) = 1.58 \text{ bits/sym}.$$

$$\text{y con } H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y),$$

$$\text{saco que } H(Y|X) = 0.66, H(X|Y) = 0.66 \text{ bits/sym}$$

5. Una fuente sin memoria tiene el alfabeto $A = \{-5, -3, -1, 0, 1, 3, 5\}$ con las probabilidades correspondientes $\{0.05, 0.1, 0.1, 0.15, 0.05, 0.25, 0.3\}$.

a. ¿Cuál es la entropía de la fuente?

$2(0.05)$ y $2(0.1)$

b. Suponga que la fuente se cuantiza de acuerdo a la siguiente regla de cuantización:

$$\begin{cases} q(-5) = q(-3) = -4 \\ q(-1) = q(0) = q(1) = 0 \\ q(3) = q(5) = 4 \end{cases}$$

¿Cuál es la entropía de la fuente cuantizada?

Calculamos la entropía igual que antes:

$$H(X) = -2 \cdot 0.05 \log_2(0.05) - 2 \cdot 0.1 \log_2(0.1) - 0.15 \log_2(0.15) - 0.25 \log_2(0.25) - 0.3 \log_2(0.3) = 2.53 \text{ bits/sym}$$

Ahora, agrupamos el alfabeto (los posibles mensajes):

El mensaje -4 agrupa a -5 y -3 con probs 0.05 y 0.1 respectivamente.

$$\text{En consecuencia: } \begin{cases} x'_1 = -4 \\ x'_2 = 0 \\ x'_3 = 4 \end{cases} A'$$

$$P(x' = -4) = P(x = -5) + P(x = -3) = 0.15$$

$$P(x' = 0) = P(x = -1) + P(x = 0) + P(x = 1) = 0.3$$

$$P(x' = 4) = P(x = 3) + P(x = 5) = 0.15$$

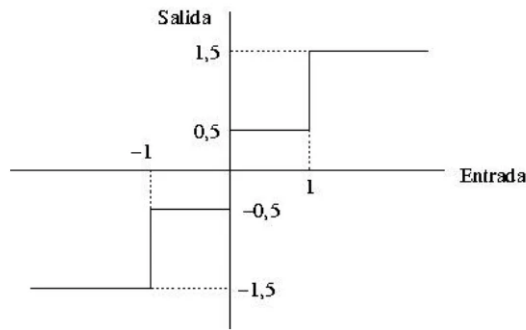
$$\left. \begin{aligned} \sum P(x' = x'_i) &= \\ &= 1 \end{aligned} \right\} \checkmark$$

$$H'(X) = -0.15 \log_2(0.15) - 0.3 \log_2(0.3) - 0.15 \log_2(0.15) =$$

$$H'(X) = 1.41 \text{ bits/sym.}$$

6. Un proceso gaussiano de media cero y varianza unidad se muestrea y las muestras se aplican a un cuantizador uniforme que tiene la característica de la figura. Calcular la entropía de la salida del cuantizador.

Atu



Tenemos que mapear las posibles salidas con sus respectivas probabilidades, que siguen una distribución Gaussiana: $\mu=0$ y $\sigma^2=1$, tengo que:

$P(S=-1.5) \Rightarrow$ área bajo la campana gaussiana hasta que

$E = -1$, es decir: *la campana es simétrica*

$$P(S=-1.5) = 1 - Q(1) = 0.1587 = P(S=1.5)$$

$\hookrightarrow$ la función Q denota la cola de probabilidad de la distribución gaussiana.

Dicho de otro modo, la probabilidad de que la salida sea -1.5 es resultado de integrar la campana de Gauss de $-\infty$ hasta -1 .

$$P(S=-0.5) = Q(1) - Q(0) = 0.3413 = P(S=0.5).$$

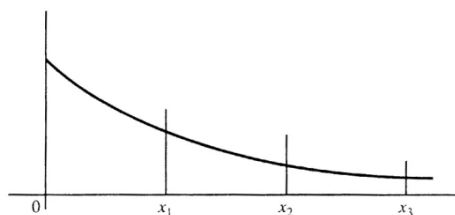
Calculamos la entropía como siempre:

$$H(x) = -2(0.3413 \log_2(0.3413)) + 0.1587 \log_2(0.1587) = 1.902 \text{ bits/sym}$$

7. La entrada a un cuantizador es una variable aleatoria que tiene una función densidad de probabilidad dada por:

$$f(x) = \begin{cases} ae^{-ax}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

La señal se va a cuantizar con un cuantizador de 4 niveles como se muestra en la figura. Determinar los valores de x_i , $i=1, 2, 3$ en términos de 'a' de forma que la entropía a la salida del cuantizador sea máxima.



Piden valores de

x_1, x_2 y x_3

Según Shannon, la entropía es máxima si todos los elementos de la señal son equiprobables a la hora de aparecer. Es decir: $P(X=0) = P(X=x_1) = P(X=x_2) = P(X=x_3) = 1/4$

Calculamos $P(X \leq x_1) = \int_0^{x_1} ae^{-ax} dx = -e^{-ax} \Big|_0^{x_1} = -e^{-ax_1} + 1$

y como $P(X \leq x_1) = 1/4 = -e^{-ax_1} + 1$, despejo que $\boxed{-e^{-ax_1} = -3/4}$,
es decir, $\ln 3/4 = -ax_1 \Rightarrow x_1 = 0.288/a$

prob $P(X \leq x_1)$

Con $P(X \leq x_2) = \int_0^{x_2} ae^{-ax} dx = -e^{-ax_2} + 1$ y $P(X \leq x_2) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

Obtenso que $x_2 = \frac{\ln 1/2}{-a} = 0.693/a$

y para $P(X \leq x_3) = (\dots) = -e^{-ax_3} + 1$ y $P(X \leq x_3) = 3/4$,

tenso que $x_3 = \frac{\ln(1/4)}{-a} = 1.386/a$

8. Una señal se muestrea 500 veces por segundo y las muestras tienen una función densidad de amplitudes gaussiana de media cero y varianza σ^2 . Las muestras se cuantizan de acuerdo a la tabla siguiente.

a) Determinar la entropía de la salida del cuantizador y la tasa de información en bits por segundo.

b) Determinar los intervalos de cuantización en función de σ de forma que a la salida del cuantizador la entropía sea máxima.

Atun

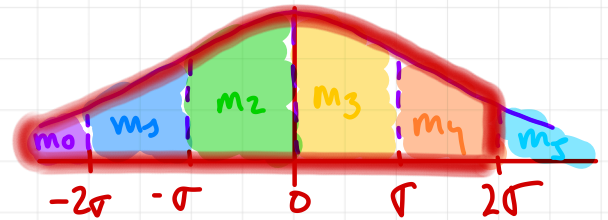
Intervalos

Entrada al cuantizador	Salida del cuantizador
$-\infty < x_i < -2\sigma$	m_0
$-2\sigma < x_i < -\sigma$	m_1
$-\sigma < x_i < 0$	m_2
$0 < x_i < \sigma$	m_3
$\sigma < x_i < 2\sigma$	m_4
$2\sigma < x_i < \infty$	m_5

$$\mu = 0$$

$$\sigma^2 = \sigma^2$$

A) Para la entropía tenemos que recurrir otra vez a la campana de Gauss para obtener las probabilidades de m_0, m_1, m_2, m_3, m_4 y m_5 .



la probabilidad de m_0 es:

$$P(m_5) = 1 - Q(2\sigma) = 1 - Q(2) = 0.0228 = P(m_0)$$

Con $Q(2\sigma)$ calculo la probabilidad hasta 2σ , pero como quiero la probabilidad de m_5 en adelante, resto 1.

Por simetría $P(m_5) = P(m_0)$.

$$P(m_4) = Q(2\sigma) - Q(\sigma) = Q(2) - Q(1) = 0.1359 = P(m_1)$$

$$P(m_3) = Q(\sigma) - Q(0) = Q(1) - Q(0) = 0.3413 = P(m_2)$$

Con las probabilidades, tengo que $H(X)$ es:

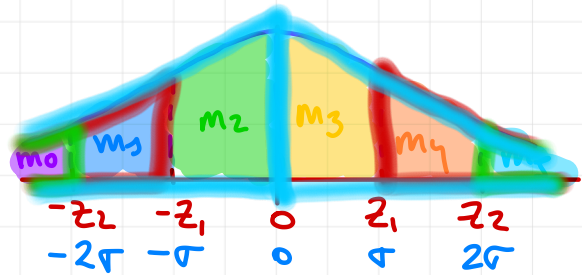
$$H(X) = 2.09 \text{ bits/sym.} \quad \text{y} \quad R(X) = 500 \text{ sym/s} \cdot H(X) = 1045 \text{ bits/s}$$

B) Entropía máxima si $P(m_0) = \dots = P(m_5) = \frac{1}{6}$.

Como la campana es simétrica, con conocer los

intervalos del lado positivo tengo automáticamente los del lado negativo.

La campana queda así:



Sé que $Q(z_2) = p(m_0) + p(m_1) + p(m_2) + p(m_3) + p(m_4) = \frac{5}{6} = 0.833$

Busco en la tabla el valor más próximo a 0.83. En este caso, $\boxed{z_2 = 0.97}$

Y con $Q(z_1) = p(m_0) + \dots + p(m_3) = \frac{4}{6} = 0.667 \Rightarrow \boxed{z_1 = 0.43}$

z_1 y z_2 son las posiciones tipificadas, es decir, $z_i = \frac{x_i - \mu}{\sigma}$.

Entonces x_1 y x_2 reales son 0.43σ y 0.97σ y la tabla queda como:

Entrada al cuantizador	Salida del cuantizador
$-\infty < x_i < -2\sigma$	m_0
$-2\sigma < x_i < -\sigma$	m_1
$-\sigma < x_i < 0$	m_2
$0 < x_i < \sigma$	m_3
$\sigma < x_i < 2\sigma$	m_4
$2\sigma < x_i < \infty$	m_5

$$m_0 \in (-\infty, -0.97\sigma)$$

$$m_1 \in (-0.97\sigma, -0.43\sigma)$$

$$m_2 \in (-0.43\sigma, 0)$$

$$m_3 \in (0, 0.43\sigma)$$

$$m_4 \in (0.43\sigma, 0.97\sigma)$$

$$m_5 \in (0.97\sigma, +\infty)$$

Atun

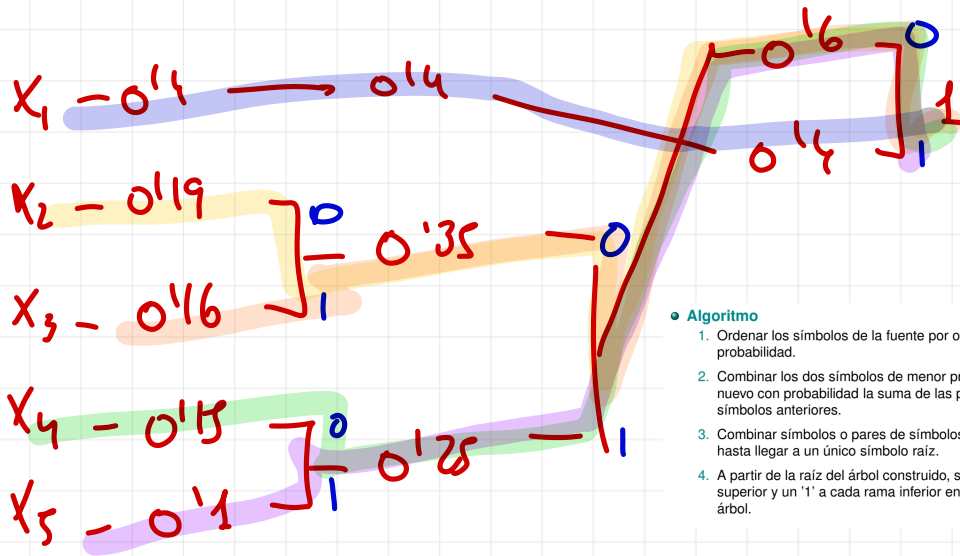
9. Una fuente DMS tiene 5 símbolos x_1, x_2, x_3, x_4 y x_5 con probabilidades $P(x_1)=0.4$, $P(x_2)=0.19$, $P(x_3)=0.16$, $P(x_4)=0.15$, $P(x_5)=0.1$. Construir un código de Huffman y calcular la eficiencia del código.

Alfaro

DMS tiene 5 símbolos = $\{x_1, \dots, x_5\}$.

$$P(x_1) = 0.4 \quad P(x_2) = 0.19$$

$$P(x_3) = 0.16 \quad P(x_4) = 0.15 \quad P(x_5) = 0.1$$



Algoritmo

1. Ordenar los símbolos de la fuente por orden decreciente de probabilidad.
2. Combinar los dos símbolos de menor probabilidad para formar uno nuevo con probabilidad la suma de las probabilidades de los símbolos anteriores.
3. Combinar símbolos o pares de símbolos de menor probabilidad hasta llegar a un único símbolo raíz.
4. A partir de la raíz del árbol construido, se asigna un '0' a cada rama superior y un '1' a cada rama inferior en cada uno de los nodos del árbol.

$$\begin{array}{ll} x_1 - 1 & x_4 - 010 \\ x_2 - 000 & x_5 - 011 \\ x_3 - 001 & \end{array} \left\{ \text{Codificación} \right.$$

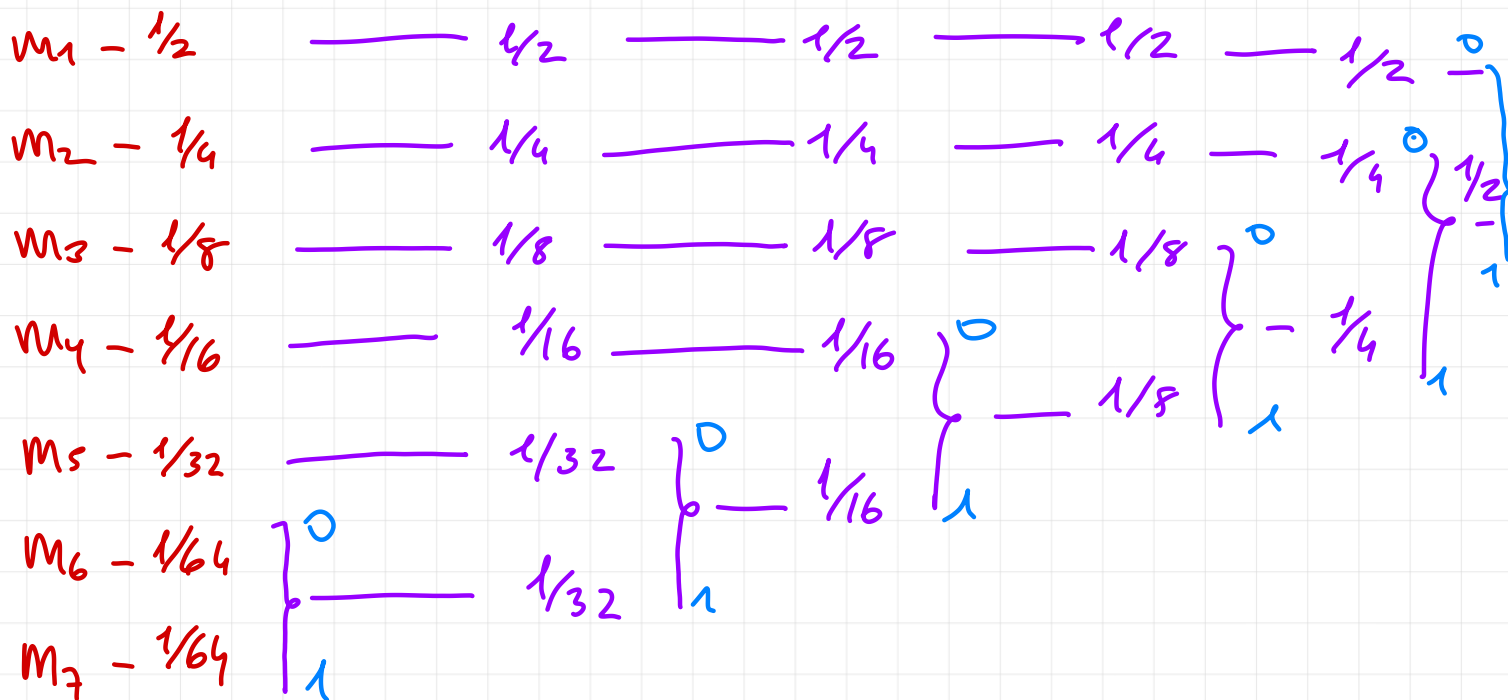
$$\bar{N} = 1 \cdot 0.4 + 3(0.16) = 2.12 \text{ bits/sym}$$

$$H = \sum_{i=1}^5 p(x_i) \cdot \log \frac{1}{p(x_i)} =$$

$$0.4 \cdot \log \frac{1}{0.4} + \dots + 0.1 \log \frac{1}{0.1} = 2.15 \text{ bits/sym}$$

$$\eta = \frac{2.15}{2.12} = 97.7\% \leftarrow \text{eficiencia}$$

- a. Encuentre la entropía de la fuente. Obtenga un código compacto binario y calcule la longitud promedio de las palabras código. Determinar la eficiencia del código.



Codificación

$$m_1 = 0$$

$$v_{m2} = 10$$

$$m_3 = 110$$

$$m_7 = 1110$$

$$m_s = 11120$$

$m_6 = 111110$

$m_7 = 1111111$

$$\bar{L} = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + 5 \cdot \frac{1}{32} + 6 \cdot \frac{1}{64} \cdot 2 = 1,969$$

$$H(X) = - \sum_i p(m_i) \log_2(m_i) = 1.969$$

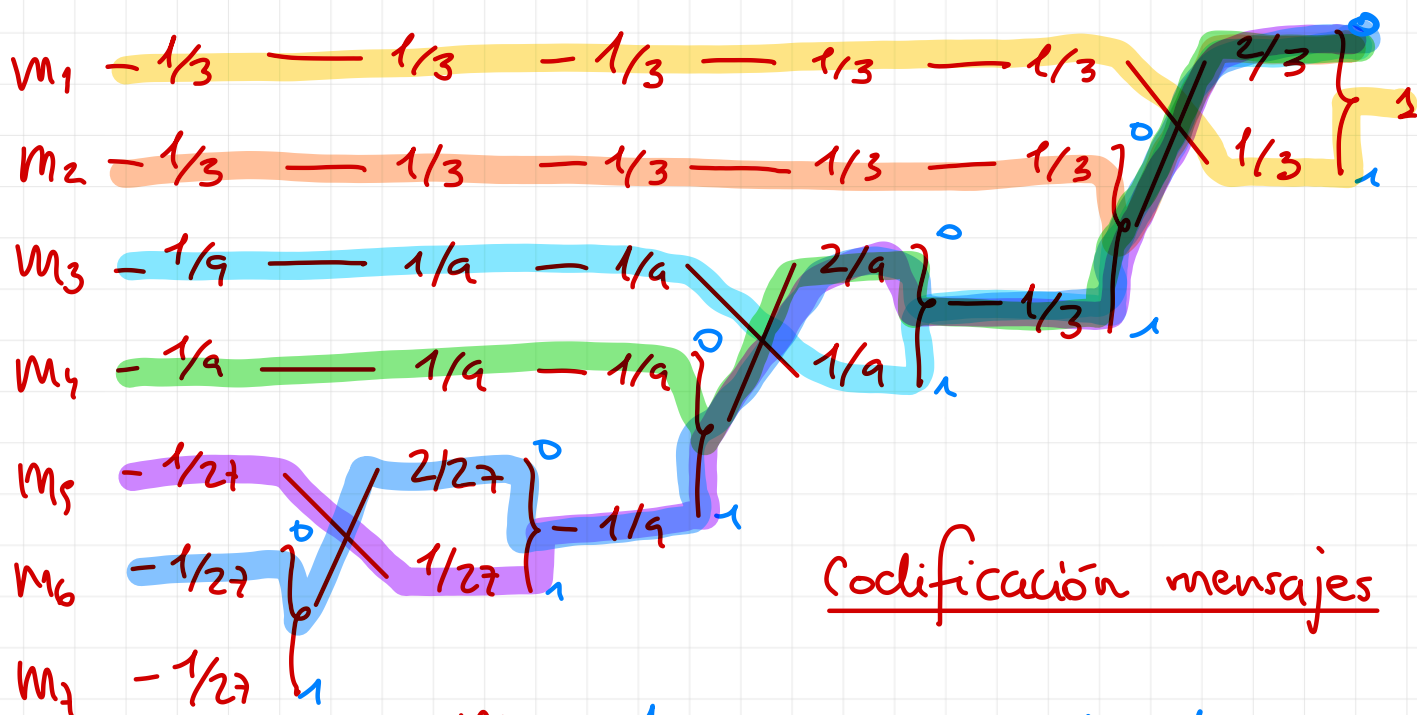
$$\eta = \frac{H(x)}{L} = \frac{1'969}{1'969} = 1 \text{ (100\%)}$$

11. Una fuente emite siete mensajes con probabilidades $1/3, 1/3, 1/9, 1/9, 1/27, 1/27$ y $1/27$, respectivamente. .

- a. Encuentre la entropía de la fuente. Obtener un código compacto binario, calcular la longitud promedio de las palabras código y determinar la eficiencia del código.

Alfaro

Igual que antes:



Codificación mensajes

$$m_1 - 1$$

$$m_5 - 01011$$

$$m_2 - 00$$

$$m_6 - 010100$$

$$m_3 - 011$$

$$m_7 - 010101$$

$$m_4 - 0100$$

$$\bar{N} = \frac{1}{3}(1+2) + \frac{1}{9}(3+4) + \frac{1}{27}(5+6+6) = 2.407 \text{ bits/sym}$$

$$H(X) = - \left(2 \cdot \frac{1}{3} \log_2 \left(\frac{1}{3} \right) + 2 \cdot \frac{1}{9} \log_2 \left(\frac{1}{9} \right) + 3 \cdot \frac{1}{27} \log_2 \left(\frac{1}{27} \right) \right) = 2.289 \text{ bits/sym.}$$

$$\eta = \frac{H(X)}{\bar{N}} = 0.951 \text{ (95.1\%)}$$

12. Una fuente emite aleatoriamente tres mensajes equiprobables e independientes.

- Encuentre la entropía de la fuente.
- Obtenga un código compacto binario y calcule la longitud promedio de las palabras código. Determinar la eficiencia del código.
- Para mejorar la eficiencia del código binario, se codifica la extensión segunda de la fuente. Obtener un código compacto binario, calcular la longitud promedio de las palabras código y determinar la eficiencia del código.

$$p(m_1) = p(m_2) = p(m_3) = 1/3$$

$$A) H(x) = -(3 \cdot 1/3 \cdot \log_2(1/3)) = -\log_2(1/3) = 1.585 \text{ bits/sym}$$

$$B) \begin{array}{l} m_1 - 1/3 \\ m_2 - 1/3 \\ m_3 - 1/3 \end{array} \left. \begin{array}{l} 2/3 \\ 1/3 \end{array} \right\} \begin{array}{l} 0 \\ 1 \end{array}$$

CODIFICACIÓN

$$m_1 - 1$$

$$m_2 - 00$$

$$m_3 - 01$$

son dos
msg. con 2 bits

$$\bar{N} = 1 \cdot 1/3 + 2 \cdot 2 \cdot 1/3 = 1.667 \text{ bits/sym}$$

$$\eta = \frac{H(x)}{\bar{N}} = 0.9508 \text{ (95.08\%)}$$

C) Básicamente tenemos que agrupar los mensajes de dos en dos. la fuente original S , y la segunda extensión S^2 .

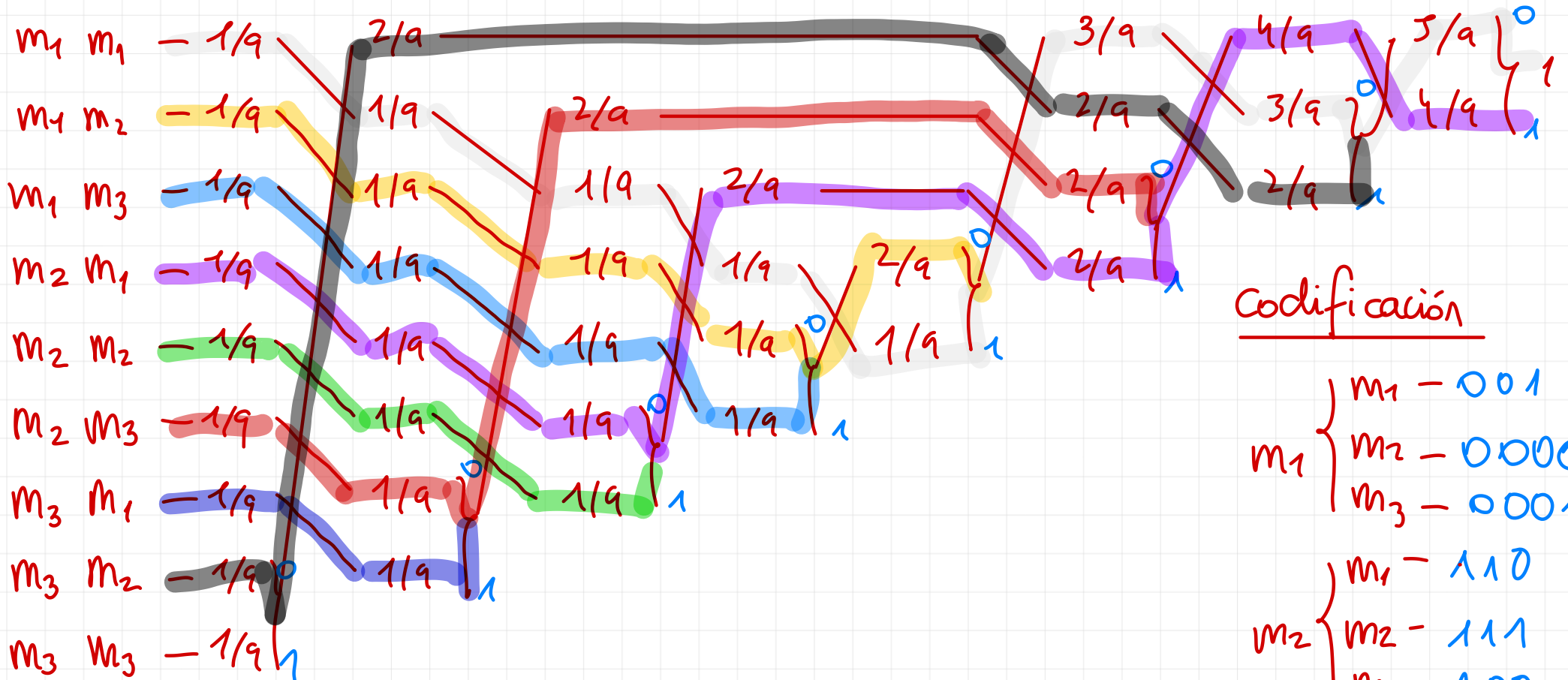
$$S^2 = \{(m_1, m_1), (m_1, m_2), (m_1, m_3), \dots, (m_3, m_2), (m_3, m_3)\}$$

Como son independientes: $p(m_i, m_j) = p(m_i)p(m_j)$.

En este caso $\forall i, j \in \{1, 2, 3\}$, $p(m_i, m_j) = 1/3 \cdot 1/3 = 1/9$.

Código compacto

Alum



Codificación

- m_1
 - $m_1 - 001$
 - $m_2 - 0000$
 - $m_3 - 0001$
- m_2
 - $m_1 - 110$
 - $m_2 - 111$
 - $m_3 - 100$
- m_3
 - $m_1 - 101$
 - $m_2 - 010$
 - $m_3 - 011$

$$\bar{N}' = 7 \cdot 3 \cdot \frac{1}{9} + 2 \cdot 4 \cdot \frac{1}{9} = 3.222 \text{ bits/sym}$$

$\uparrow$ de 3 bits $\uparrow$ de 4 bits

$$H'(X) = -9 \cdot \frac{1}{9} \cdot \log_2\left(\frac{1}{9}\right) = -\log_2\left(\frac{1}{9}\right) = 3.170 \text{ bits/sym}$$

$$\eta' = 0.984 \quad (98.40\%)$$

13. Un alfabeto de entrada consta de 100 caracteres.

- a. Si cada carácter se codifica con un código de longitud fija, determinar el número necesario de bits para codificarlos.
- b. Vamos a hacer la suposición de que 10 caracteres son igualmente probables y que cada uno ocurre con una probabilidad de 0.05. Vamos a suponer que el resto de los 90 caracteres son igualmente probables. Determinar el número de bits necesarios para codificar este alfabeto utilizando un código Huffman de longitud variable.

primeros 10 caracteres
prob 50%
el resto, 50%.

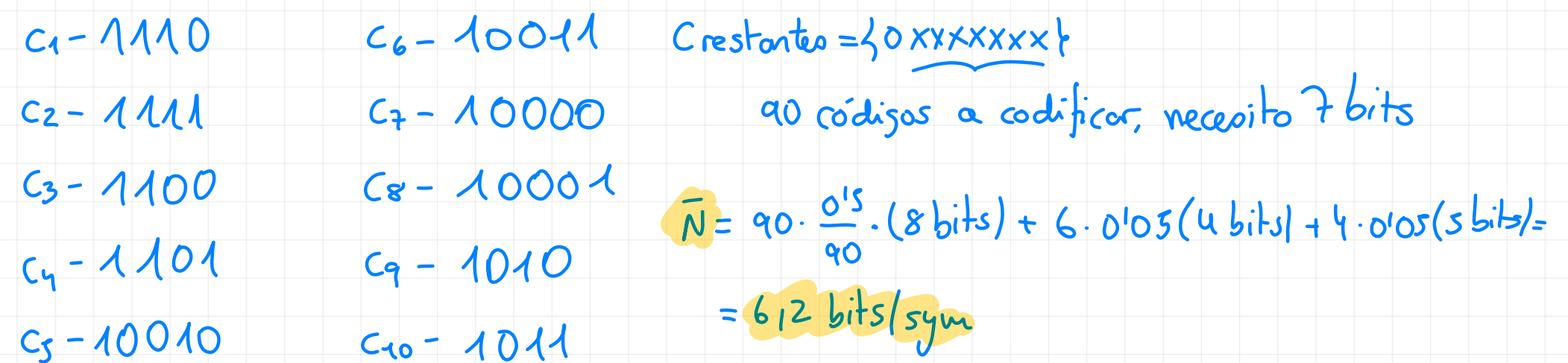
A) 100 caracteres de longitud fija:

busco n tal que $2^n \geq 100$ caracteres:

$\log_2 100 = n = 6,64$. (pero n es natural, entonces $n=6$).

Compruebo que $2^6 \geq 100$; $64 \not\geq 100$. Como no son caracteres suficientes, le sumo 1 a n : $2^7 \geq 100$; $128 \geq 100$.

Es decir, necesitamos mensajes de 7 bits.



Restantes = $\{0 \underbrace{xxxxxxx}\}$

90 códigos a codificar, necesito 7 bits

$$\bar{N} = 90 \cdot \frac{0.5}{90} \cdot (8 \text{ bits}) + 6 \cdot 0.05 (4 \text{ bits}) + 4 \cdot 0.05 (5 \text{ bits}) = 6.2 \text{ bits/sym}$$