

# Teoría de la información y la codificación

15 de octubre de 2019

---

## 1. Ejercicio 1

Sea una fuente  $S$  con un alfabeto  $A = \{a, b, c, d, e, f\}$ , con probabilidades de envío de cada mensaje  $p(a) = 0.3$ ,  $p(b) = 0.25$ ,  $p(c) = 0.2$ ,  $p(d) = 0.1$ ,  $p(e) = 0.1$ ,  $p(f) = 0.05$ , que se desea codificar.

1. Explique qué es la esperanza de información de la fuente y calcule la esperanza de información de la fuente  $S$ .

**Respuesta:**

Podemos definir esperanza de la información de una fuente o entropía a la cantidad de información promedio que contienen los símbolos usados. Y viene expresada por la siguiente formula:

$$\begin{aligned} E\{I(S)\} &= p(S = s_1) * I(S = s_1) + p(S = s_2) * I(S = s_2) + \dots + p(S = s_n) * I(S = s_n) = \\ &= \sum_{i=1}^n p(S = s_i) * I(S = s_i) \end{aligned}$$

La esperanza de información de la fuente  $S$  sera:

$$\begin{aligned} E\{I(S)\} &= -\log_6 0.3 - \log_6 0.25 - \log_6 0.2 - \log_6 0.1 - \log_6 0.1 - \log_6 0.05 \\ &= -\log_6 0.0000075 = 6.58604 \end{aligned}$$

2. Explique qué es un código descifable sin hacer uso de las explicaciones textuales de las diapositivas de clase. Ponga ejemplos de códigos descifrables y códigos indescifrables, justificando cada uno.

**Respuesta:**

Un código descifable es cuando existe para cualquier elemento del alfabeto de la fuente, una vez codificado con algún método/ algoritmo/ función, un único elemento del alfabeto del receptor. Y ademas existe algún método/ algoritmo/ función (que sera la inversa de la función de codificación) para volver a tener el elemento original de forma única.

*Ejemplo código descifable:*

$$\begin{aligned}
f &: S \rightarrow C \\
X \in S, C \in C, X &\rightarrow f(X) \\
S = \{a, b, \dots, z\}, A &= \{0, 1\}, C = \{00000, 00001, \dots, 11100\} \\
a &\rightarrow f(a) = 00000 \\
b &\rightarrow f(a) = 00001 \\
&\dots \\
z &\rightarrow f(a) = 11100
\end{aligned}$$

Como podemos ver tenemos un código descifrable. Esto se debe a que nuestra función  $f(x)$  es biyectiva. Esto es, como hemos explicado intuitivamente arriba, que para elemento de S existe un único elemento de C y viceversa, que para cada elemento de C existe un único elemento de S.

*Ejemplo código indescifrable:*

$$\begin{aligned}
f &: S \rightarrow C \\
X \in S, C \in C, X &\rightarrow f(X) \\
S = \{a, b, \dots, z\}, A &= \{0, 1\}, C = \{0, 1, \dots, 0\} \\
a &\rightarrow f(a) = 0 \\
b &\rightarrow f(b) = 1 \\
&\dots \\
z &\rightarrow f(z) = 0
\end{aligned}$$

Como podemos ver este código es indescifrable ya que  $f(x)$  no es inyectiva. Esto significa que para algún elemento de S hay mas de un elemento en C. Por ejemplo a la letra **a** y **z** les corresponde a los dos el **0**.

3. Explique qué es un código uniforme, y diseñe un código ternario uniforme en el alfabeto  $B = \{0, 1, 2\}$ , con el mínimo número de longitud de palabra posible que permita codificar todos los símbolos del alfabeto de la fuente. Explique si el código es descifrable o no.

**Respuesta:**

Un código uniforme es aquel en el que todas las palabras de  $A^*$  que existen en  $C$  tienen la misma longitud. En definitiva que todos los elementos de C tengan el mismo número de símbolos.

*Código ternario uniforme no descifrable:*

Si no nos importa si el código es descifrable o no, podemos realizar un código uniforme de longitud uno. Pero de esta manera la función no sería inyectiva y por lo tanto indescifrable.

$$\begin{aligned}
h &: S \rightarrow C \\
X \in S, C \in C, X &\rightarrow h(X) \\
B = \{0, 1, 2\} A = \{a, b, c, d, e, f\}, C &= \{0, 1, 2, 0, 1, 2\} \\
a &\rightarrow h(a) = 0 \\
b &\rightarrow h(b) = 1 \\
c &\rightarrow h(c) = 2 \\
d &\rightarrow h(d) = 0 \\
e &\rightarrow h(e) = 1 \\
f &\rightarrow h(f) = 2
\end{aligned}$$

*Código ternario uniforme descifrable:*

El anterior tiene el mínimo número de símbolos pero es indescifrable, para realizarlo

descifrable debería ser el siguiente:

$$\begin{aligned}
 h : S &\rightarrow C \\
 X \in S, C &\in C, X \rightarrow h(X) \\
 B = \{0, 1, 2\} \quad A &= \{a, b, c, d, e, f\}, \quad C = \{00, 01, 02, 10, 11, 12\} \\
 a &\rightarrow h(a) = 00 \\
 b &\rightarrow h(b) = 01 \\
 c &\rightarrow h(c) = 02 \\
 d &\rightarrow h(d) = 10 \\
 e &\rightarrow h(e) = 11 \\
 f &\rightarrow h(f) = 12
 \end{aligned}$$

Estamos usando el mínimo numero de longitud de palabra posible ya que si usáramos longitud 1 solo podríamos codificar 3 elementos y que el código fuera descifrable. En cambio con longitud dos podemos llegar a codificar  $3^2 = 9$  símbolos del alfabeto de la fuente, aunque en este caso solo usemos seis.

4. Diseñe un código binario en el alfabeto  $B=0,1$ , instantáneo no uniforme, que codifique los 6 símbolos. Se valorará que el código sea óptimo (no exista otro código con una longitud de palabra promedio menor).

**Respuesta:**

$$\begin{aligned}
 h : S &\rightarrow C \\
 X \in S, C &\in C, X \rightarrow h(X) \\
 B = \{0, 1\} \quad A &= \{a, b, c, d, e, f\}, \quad C = \{0, 100, 101, 110, 1110, 1111\} \\
 a &\rightarrow h(a) = 0 \\
 b &\rightarrow h(b) = 100 \\
 c &\rightarrow h(c) = 101 \\
 d &\rightarrow h(d) = 110 \\
 e &\rightarrow h(e) = 1110 \\
 f &\rightarrow h(f) = 1111
 \end{aligned}$$

Puede ser que exista un código instantáneo mejor, pero no se me ocurre como realizarlo.

5. Indique la diferencia entre código instantáneo y código de traducción única. Diseñe un código binario en el alfabeto  $B=0,1$ , que sea de traducción única que no sea instantáneo, que codifique los 6 símbolos. Explique el motivo por el que es de traducción única, pero no instantáneo.

**Respuesta:**

Los códigos instantáneos son un subconjunto de los códigos de traducción única. Cuando un código es instantáneo significa que ninguna palabra codificada coincide con el comienzo de otra. En cambio los códigos de traducción única son cuando para cualquier sucesión única de palabras de códigos a transmitir, corresponde una única sucesión de símbolos

transmitidos.

$$\begin{aligned}
 h : S &\rightarrow C \\
 X \in S, C &\in C, X \rightarrow h(X) \\
 B = \{0, 1\} &A = \{a, b, c, d, e, f\}, C = \{1, 101, 1001, 10001, 100001, 1000001\} \\
 a &\rightarrow h(a) = 1 \\
 b &\rightarrow h(b) = 101 \\
 c &\rightarrow h(c) = 1001 \\
 d &\rightarrow h(d) = 10001 \\
 e &\rightarrow h(e) = 100001 \\
 f &\rightarrow h(f) = 1000001
 \end{aligned}$$

El código como se puede ver es de tradición única ya que para cualquier sucesión de símbolos que recibamos existe una única sucesión de palabras posible y no es instantáneo ya que por ejemplo todas las palabras comienzan por la palabra **a** por ejemplo.