TEMA 9
TEORÍA DE LA INFORMACION

BIBLIOGRAFÍA
[Abramson 80]

Norman Abramson,
"Teoría de la información y codificación",
Ed. Paraninfo.

Introducción

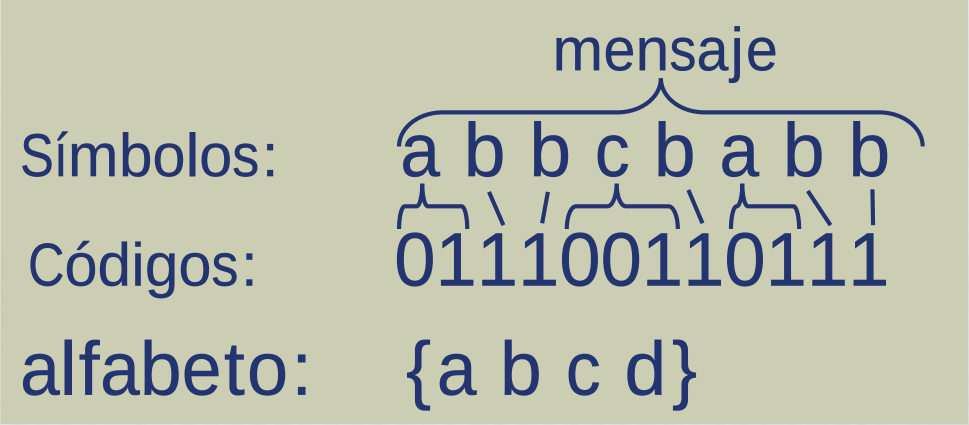
Claude E. Shannon 1948 (Bell) “Una teoría matemática de la comunicación”.

- Fuentes de información.
- Información y codificación.
- Medidas relacionadas con la información.
- Teorema de Shannon.

Claude Elwood Shannon (1916 – 2001)
Conocido como “el padre de la teoría de la información”

Terminología

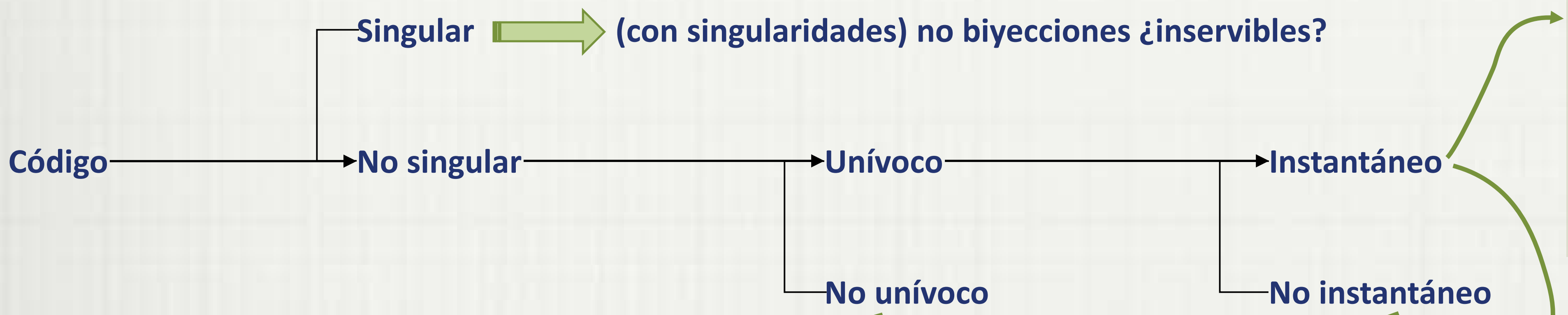
- La información se contiene en mensajes.
- Un emisor de mensajes se conoce como fuente de información.
- Los mensajes son cadenas de símbolos.
- El conjunto de símbolos utilizados para componer mensajes se denomina alfabeto.
- Los mensajes se codifican estableciendo un código para cada símbolo.



Codificación

Algunas observaciones sobre codificación utilizando un alfabeto binario.

- Necesidad de reversibilidad del código
- Optimización en función de la naturaleza de la información



Símbolo	código
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001

Símbolos del mensaje	Palabras código
s1	0
s2	01
s3	001
s4	111

111001 → s4 s1 s2 s4 s3

Símbolos del mensaje	Palabras código
s1	1100
s2	110
s3	10

1 0 1 1 0 1 0

s3 s2 s3

Símbolos del mensaje	Palabras código
s1	1110
s2	110
s3	10

1 0 1 1 0 1 0

s3 s2 s3

Otro tema diferente, pero relacionado es la robustez frente a errores: códigos bloque… Reed–Solomon, Hamming, Hadamard, Justesen, Golay, Reed–Muller, grafos de expansion bipartidos,…

Longitud media de una codificación

¿Cual es el código más escueto?… Binits (Binary Digits) de una codificación.

Mensaje	Símbolo	Probabilidad	Código A	Código B
Sol	s1	1/10	00	1110
Nubes	s2	3/10	01	10
Sirimiri	s3	5/10	10	0
Tormenta	s4	1/10	11	110

Mensaje (tiempo de la última semana):

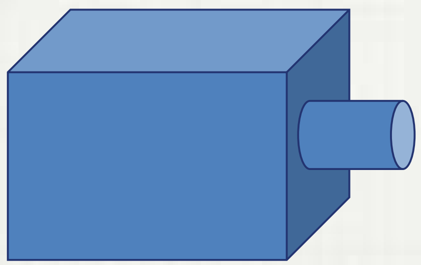
	s1	s2	s3	s3	s4	s3	s3
Código A:	00	01	10	10	11	10	10
Código B:	1110	10	0	0	110	0	0

L(A) = 2 binits/símbolo

L(B) = 4 (1/10) + 2 (3/10) + 1 (5/10) + 3 (1/10) =1,8 binits/símbolo

Volver a la página 13 del tema 3, pdf 2

Fuentes de Markov (procesos markovianos)

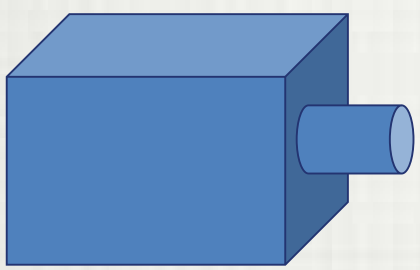


... a b b c b a b b

P(a)=2/8
P(b)=4/8
P(c)=1/8
P(d)=1/8

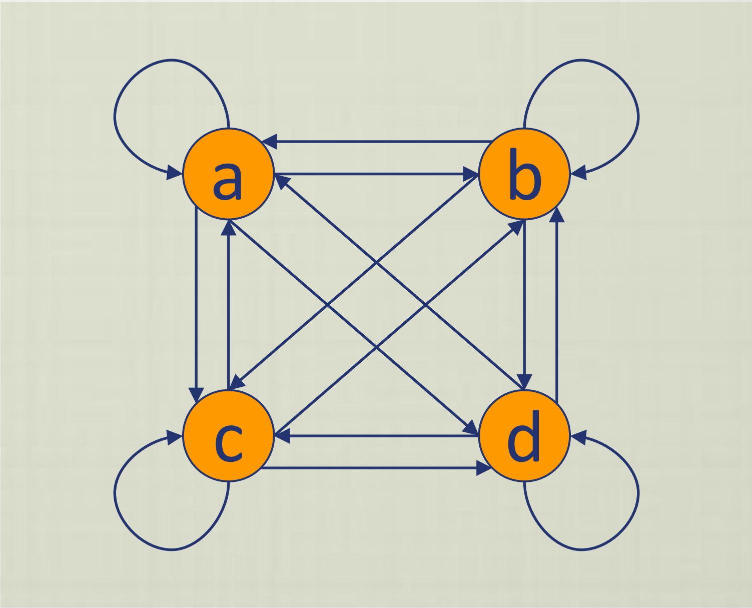
Fuente de memoria nula

Fuente markoviana de orden 1



... a b b c b a b b

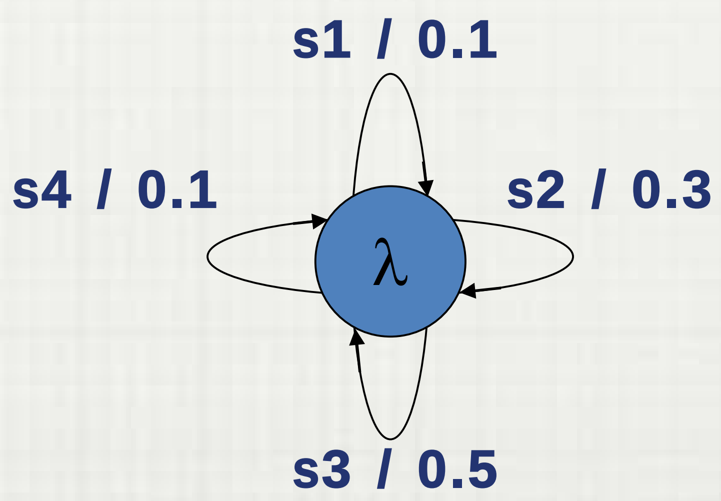
P(a|a)	**P(a|b)**	**P(a|c)**	**P(a|d)**
P(b|a)	**P(b|b)**	**P(b|c)**	**P(b|d)**
P(c|a)	**P(c|b)**	**P(c|c)**	**P(c|d)**
P(d|a)	**P(d|b)**	**P(d|c)**	**P(d|d)**



Cadena Markoviana: modelo estocástico dependiente de una variable de estado.
Orden 1: la variable de estado depende únicamente del símbolo anterior

Una Fuente

Mensaje	Símbolo	Probabilidad
Sol	s1	1/10
Nubes	s2	3/10
Sirimiri	s3	5/10
Tormenta	s4	1/10



Mensaje (tiempo de la última semana):
s1 s2 s3 s3 s4 s3 s3

Cantidad de información.

Bits (Binary units), Hartleys*, Nats y otros.

$$^{*}I(X) = \log_a\left(\frac{1}{P(X)}\right)$$

Mensaje	Probabilidad
Sol	1/10

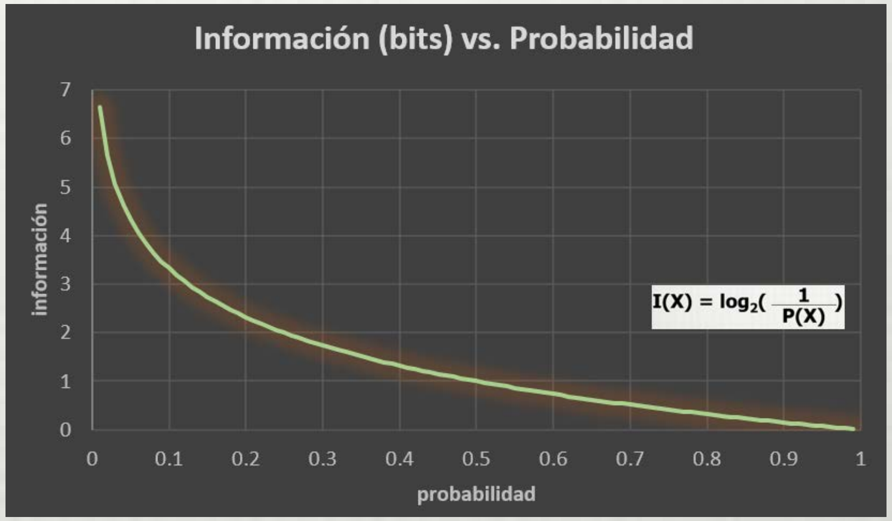
$$I(Sol) = \log_2\left(\frac{1}{P(Sol)}\right) = \log_2(10) = 3.32 \text{ bits}$$

$$I(Sol) = \ln\left(\frac{1}{P(Sol)}\right) = \ln(10) = 2.30 \text{ nats}$$

$$I(Sol) = \log_{10}\left(\frac{1}{P(Sol)}\right) = \log_{10}(10) = 1 \text{ Hartley}$$

1 bit es la cantidad de información de un evento (mensaje, símbolo, binit,...) para el que es equiprobable su aparición o no (es decir, P(X)=0.5).

$$P(X)=0.5,\ \ I(X) = \log_2\left(\frac{1}{P(X)}\right) = \log_2(2) = 1 \text{ bit}$$



Entropía de una fuente.

Entropía = valor medio de información por símbolo

$$\boxed{H(S) = \sum_i P(s_i)\ I(s_i) = \sum_i P(s_i) \log_2 \frac{1}{P(s_i)}}$$

Mensaje	Símbolo	Probabilidad
Sol	s1	1/10
Nubes	s2	3/10
Sirimiri	s3	5/10
Tormenta	s4	1/10

$$H(S) = 0.1 \log_2(10) + 0.3 \log_2(10/3) + 0.5 \log_2(2) + 0.1 \log_2(10) = 1{,}68 \text{ bits/símbolo}$$

Este es un ejemplo para una fuente de orden cero. Para fuentes de orden superior, no se trata de un cálculo trivial. Más adelante se amplia un poco este tema.

Cálculo de la entropía

Al definir la entropía se vio un ejemplo con una fuente de memoria nula. Para realizar el cálculo con una fuente de orden mayor se debe calcular la entropía de cada estado de igual modo al visto y después obtener la media global teniendo en cuenta la probabilidad de los estados. Esto solo tiene sentido cuando existe una probabilidad estacionaria, cosa que no siempre sucede.

El estado estacionario exige que la distribución de probabilidades no varíe de un instante de tiempo al siguiente:

$$[p(s_0) \ldots p(s_n)]_{t+1} = [p(s_0) \ldots p(s_n)]_t \, [P_{ij}] \qquad (1)$$

Si $[p(s_0) \ldots p(s_n)]_{t+1} = [p(s_0) \ldots p(s_n)]_t$

$$[p(s_0) \ldots p(s_n)] \, ([I]-[P_{ij}])=0$$

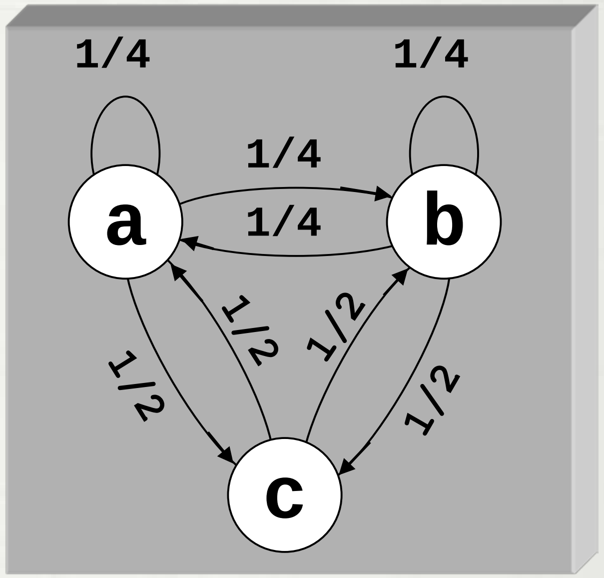
En general, las probabilidad del vector de estados en el instante t puede obtenerse en función de la probabilidad en el estado inicial mediante una expresión que puede derivarse de (1) mediante la aplicación de la transformada z.

Cálculo de la entropía (con estados)

Enunciado: se tiran al aire dos monedas. Con dos caras se emite "a", dos cruces "b" y en otro caso "c". Si se emite "c", la siguiente tirada se emite "a" si iguales y "b" si distintas, y se continua con la primera interpretación

CC a
++ b
C+ } c
+C } c

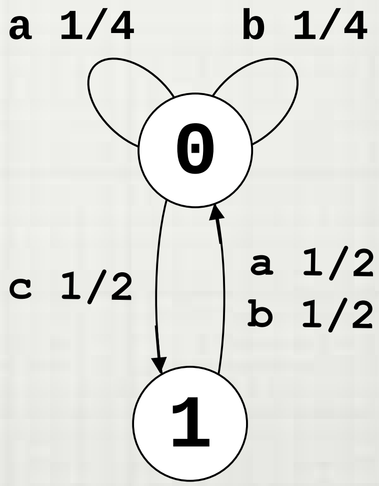
CC } a
++ } a
C+ } b
+C } b



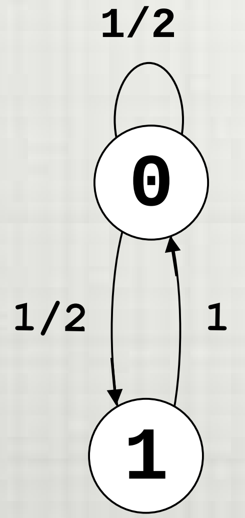
$H_0(S) = 1/4 \log_2 4 + 1/4 \log_2 4 + 1/2 \log_2 2 = 3/2$
$H_1(S) = 1/2 \log_2 2 + 1/2 \log_2 2 = 1$

$H(S) = P_0 H_0(S) + P_1 H_1(S) = 2/3\ 3/2 + 1/3\ 1 = 1.33$

simplificación



Cálculo de la probabilidad de los estados estacionarios
--condición: MM regular (irreductible, recurrente, aperiódico)--

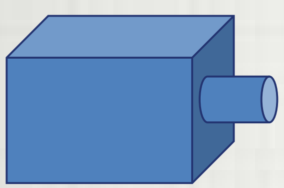


$$(P_0\ P_1) = (P_0\ P_1)\begin{pmatrix} 1/2 & 1/2 \\ 1 & 0 \end{pmatrix}$$

$$\left.\begin{array}{l} \left.\begin{array}{l} P_0 = 1/2\, P_0 + P_1 \\ P_1 = 1/2\, P_0 \end{array}\right\} = \\ \\ P_1 + P_0 = 1 \end{array}\right\} \begin{array}{l} P_0 = 2/3 \\ P_1 = 1/3 \end{array}$$

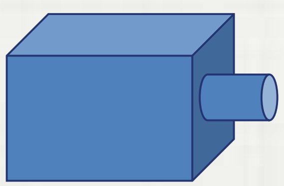
Extensión de orden N de una fuente

Se denomina extensión de orden n de una fuente dada, a otra con emisiones semejantes donde cadenas de longitud n son substituidas por nuevos símbolos



... a b b c b a b b

P(a)=2/8
P(b)=4/8
P(c)=1/8
P(d)=1/8



... s_2 s_7 s_5 s_6

aa	ab	ac	ad	ba	bb	bc	bd
s_1	**s_2**	**s_3**	**s_4**	**s_5**	**s_6**	**s_7**	**s_8**
4/64	8/64	2/64	2/64	8/64	16/64	4/64	4/64

ca	cb	cc	cd	da	db	dc	dd
s_9	**s_{10}**	**s_{11}**	**s_{12}**	**s_{13}**	**s_{14}**	**s_{15}**	**s_{16}**
2/64	4/64	1/64	1/64	2/64	4/64	1/64	1/64

¡OJO!: este ejemplo supone que la fuente es ciertamente de memoria nula

Primer Teorema de Shannon

Se conoce también como "Teorema de la codificación sin ruido"

"Siempre" puede obtenerse una codificación cuya longitud media coincida con la entropía de la fuente.

$$\boxed{\lim_{n \to \infty} \frac{L_n}{n} = H_r(S)}$$

(r es la base de representación utilizada)

Obsérvese que:

El valor de la entropía marca un límite a la compresión de mensajes (su codificación)

Puede aproximarse al límite cuanto se quiera recurriendo a extensiones de orden de la fuente.

(No se confunda el código límite con código compacto:

el de menor longitud media para una codificación de cada símbolo original de la fuente.)

Ejercicio

Se establecen varios códigos para una fuente según la tabla. [Abramson 80, p. 79]

¿Cual de los códigos es unívocamente decodificable?

¿Cual es instantáneo?

Calcular la longitud media de todos los códigos unívocos.

Calcular la entropía de la fuente.

Símbolo	Probabilidad	A	B	C	D	E	F
s1	1/2	000	0	0	0	0	0
s2	1/4	001	01	10	10	10	100
s3	1/16	010	011	110	110	1100	101
s4	1/16	011	0111	1110	1110	1101	110
s5	1/16	100	01111	11110	1011	1110	111
s6	1/16	101	011111	111110	1101	1111	001
		✓	✓	✓	✗ 10110	✓	✗ 00100
		✓	✗	✓	-	✓	-
		3	2.125	2.125	1.9375	2	2

$\boxed{H=2}$

Revisión

Hartley, Shannon, Markov.
Fuente de Información, mensaje, símbolo, alfabeto, código.
Fuente de Markov de orden n (para n=0, Fuente de memoria nula).
Extensión de orden m de una fuente.
Códigos
    (no)singular, unívoco, instantáneo, compacto.
Longitud media de una codificación - binits.
Información (medida)
    bits, nats, hartleys.
Entropía
PRIMER TEOREMA DE SHANNON
Fuente ergódica /no ergódica