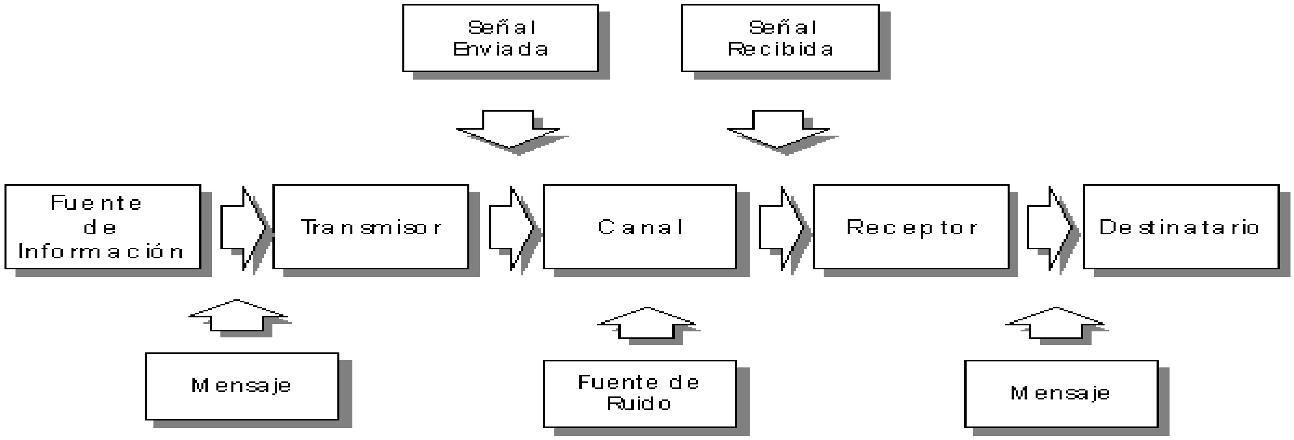
Teoría de la Información - Modelo de comunicación



¿Cuánta información transmite la fuente?

Notar: Emisor y Receptor deben tener un alfabeto común.

La información

"Definimos como información a todas aquellas representaciones simbólicas que por el significado que le asigna quien la recibe e interpreta, contribuyen a disminuir la incertidumbre"

- **Un suceso contendrá mayor cantidad de información, a menor probabilidad de ocurrencia.**
- **El concepto de información supone la existencia de duda o incertidumbre.**

EFECTOS DE LA INFORMACIÓN (Shannon):

- **"Información es lo que reduce la incertidumbre"**
- **"Es el conjunto de datos que permiten tomar una decisión".**

La jerarquía de la información

- **Datos: fragmentos inconexos** ***(140 km/*h).**
- **Información: datos organizados o procesados (140 km/h, circulando por autopista).**
- **Conocimiento: es información internalizada** ***(velocidad* excesiva).**
- **Sabiduría: es conocimiento integrado** ***(si circulo a* velocidad excesiva = multa).**

Medida de la Información

El valor

La Información es lo incomprensible **que posee un mensaje, es lo** original **que contiene**

- **Existirá Mensaje: sólo si contiene, o es portador de información.**
- **Mensaje de información máxima: Si sus elementos son originales al 100% : es ininteligible**

El VALOR de un mensaje es mayor si es más nuevo (aporta mas al aprendizaje del receptor)

<u>Medir</u> la cantidad de información de un mensaje **es medir lo** IMPREVISIBLE **que contiene .**

Equivale a medir la PROBABILIDAD **de que algo sea imprevisible para el receptor.**

Redundancia

Al formar una palabra se elige una primera letra de todas las posibles; luego, se elige la segunda letra cuya probabilidad depende de la primera letra seleccionada, y así hasta formar la palabra. La información aportada por las últimas letras de una palabra es menor.

En los idiomas permite que si se pierde parte de un mensaje podamos completarlo.

Pr jmpl, qtms tds ls vcls.

La redundancia es que nos permite ahorrar tiempo en la decodificación de los mensajes. Generalmente, no leemos cada una de las letras y palabras que conforman un texto, sino que vamos adivinando lo que viene:

Sgeún un etsduio de una uivenrsdiad ignlsea, no ipmotra el odren en el que las ltears etsan ersciats, la uicna csoa ipormtnate es que la pmrirea y la útlima ltera esten ecsritas en la psioción cocrrtea. El rsteo peuden estar ttaolmntee mal y aun pordas lerelo sin pobrleams.

Principios de la medición de información

Primer Principio:

- **La información que suministra un evento es función de la inversa de probabilidad de ocurrencia del mismo:**

$$I = f(1/P)$$

Es <u>la probabilidad que tiene un mensaje de ser enviado</u> (no su contenido), <u>lo que determina su valor informativo</u>.

Segundo Principio (ó información mutua):

- **La información total recibida será la suma de las informaciones mutuas.**

Si ocurren dos eventos simultáneamente: ***a*** *y* ***b****, que nos aportan información:*

$$I = I_a + I_b = f(1/P_a) + f(1/P_b)$$

Pero la información deberá ser esa misma función de a y b simultáneamente, es decir:

$$I = f(\ 1 / \ P_{a \ \underline{y}\ b}\ )$$

Si los eventos son independientes, la probabilidad de que ocurran simultáneamente es:

$$P(a \text{ y } b) = P(a)*P(b)$$

De (1) y (2) vemos que las informaciones se suman, mientras que las probabilidades se multiplican.

La función que nos permite realizar esto es evidentemente una función logarítmica, o sea:

$$I = \log(1/P)$$

UNIDADES DE INFORMACION

$I(a) = \log_2 \frac{1}{P(a)}$ BIT (ó también: SHANNON)	
$I(a) = \log_e \frac{1}{P(a)}$ (NAT)	$I(a) = \log_{10} \frac{1}{P(a)}$ (HARTLEY)

BIT Y BINITS

Señales Binarias con dos niveles de tensión. Al "0" y al "1" se les llama dígitos binarios.

Si cada dígito es equiprobable: P(0) = P(1) = 1/2, la información por cada dígito binario será:

$$I = \log(1/½) = \log 2 = 1 \text{ [bit]}$$

- **A un dígito binario se lo llama *binit***
- **A la unidad de información se la llama *BIT*.**

<u>BIT es la unidad de información que transporta un dígito binario.</u>

(No siempre un dígito binario transporta un bit de información)

- **Si los dígitos binarios son conocidos de antemano (010101010101)**
- **Si los ceros y los unos *no son equiprobables***

Dígitos binarios que CONTIENEN información:

- **Forman códigos, representan letras, números y símbolos especiales.**

Dígitos binarios que NO CONTIENEN información:

- **Bits de paridad, bits de sincronismo para principio y fin de transmisiones.**

SIMBOLOS Y DATOS

Símbolo: Todo aquello que por una convención predeterminada hace alusión a algo que <u>no necesariamente debe estar presente</u>.

Atributos: Son las *propiedades o cualidades* de los sucesos al representarse simbólicamente.

Datos: Son *representaciones simbólicas de propiedades o cualidades de entes o sucesos*.

Información de un símbolo:

Si una fuente sin memoria que entrega símbolos:

- **Antes de la aparición del símbolo: estado de incertidumbre, desconocimiento del símbolo que aparecerá.**
- **En la aparición del símbolo: sorpresa, porque aparece un símbolo no esperado.**
- **Tras la aparición del símbolo: aumento en la información que tenemos ya que no tenemos la incertidumbre anterior.**

$$I(s_i) = \log_2\left(\frac{1}{p_i}\right) \qquad \textbf{(Bits)}$$

La información de un símbolo tiene las siguientes propiedades:

- $I(s_i) \geq 0$, **(probabilidad: entre 0 y 1). El símbolo nunca supondrá pérdida de información.**
- $I(s_i) = 0 \Leftrightarrow p_i = 1$. **Si sabemos que símbolo va a aparecer, no aporta información.**
- $I(s_i) < I(s_j) \quad \text{para} \quad p_i > p_j$, **(a mayor sorpresa, mayor información).**
- $I(s_i s_j) = I(s_i) + I(s_j)$. **La información aportada por un símbolo que es la concatenación de otros dos es la suma de las informaciones de ambos símbolos.**

Fuentes de información

- **Una fuente de información es un elemento que entrega información**
  - **Fuentes sin memoria: los símbolos que emite son estadísticamente independientes.**
  - **Fuentes con memoria: la aparición de los símbolos no es estadísticamente independiente.**

Entropia

Magnitud que mide la información contenida en un flujo de datos con "m" simbolos diferentes.

La entropía "H" de una fuente es igual a la cantidad de información <u>media</u> de sus mensajes.

$$H = I_{med}$$

$$H = p_1 * \log(1/p_1) + p_2 * \log(1/p_2) + \ldots + p_m * \log(1/p_m)$$

- **Si en los mensajes sus probabilidades son iguales, la entropía total será: $H = \log_2 m$**
- **La entropía <u>nos indica el límite teórico para la compresión de datos</u>.**
- **Se utiliza habitualmente el logaritmo en base 2, y entonces la entropía se mide en <u>bits</u>.**

Ejemplo :

Transmitamos mensajes del abecedario. (combinaciones aleatorias y mensajes equiprobables):

  - **26 Letras + 5 signos de puntuación + 1 espacio en blanco = 32 símbolos**

La entropía será: $H = \log_2 32 = 5$

- **Significa que se necesitan 5 bits para codificar cada símbolo: 00000, 00001, 00010, etc.,**
- **La entropía nos permite ver la cantidad de bits necesarios para representar el mensaje que se va a transmitir.**
- **Un mensaje suministra la mayor cantidad de información cuando todos los elementos del mensaje son equiprobables.**

Información de una fuente.

- **Si no todos los binits pueden traer un bit de información.**
- **Si la información mutua depende de los símbolos que una fuente puede producir.**
- **Si el diseño de las fuentes es para todo tipo de mensajes y no para un mensaje particular.**

Entonces, se debe definir la fuente en términos de *información promedio* **o** *entropía de la fuente***.**

Entropía de fuentes sin memoria

- **La entropía H, representa la "incertidumbre media" en la ocurrencia de cada símbolo.**
- **Permite describir la fuente en términos de la información promedio producida por ella.**
- **Se puede tener el caso de que varias entropías sean iguales pero una fuente más rápida (entrega más símbolos por unidad de tiempo) que otra.**

La información que entregue la fuente será el valor medio de las informaciones que entregue cada símbolo individualmente **cada vez que aparezcan. Este parámetro se llama** <u>*Entropía*</u> **de la fuente, y se puede expresar como:**

$$\mathrm{H(S)} = E\{I(s_m)\} = \sum_{i=1}^{m} p_i \cdot I(s_i)$$

(E{·} es la esperanza matemática.)

La entropía mide la información media, y por tanto, la cantidad media de símbolos necesarios.

Propiedades de H(S):

- $0 \le \mathrm{H(S)} \le \log_2(m)$**, no es negativa y está acotada superiormente. La fuente no puede suponer una pérdida de información. No puede entregar información ilimitada.**
- $\mathrm{H(S)} = 0 \Leftrightarrow p_i = 1$ **para algún *i*. En este caso el resto de las probabilidades serán nulas. Entropía nula (caso de menor incertidumbre).**
- $\mathrm{H(S)} = \log_2(m) \Leftrightarrow p_i = 1/m \ \forall i$ **.Si todos los símbolos son equiprobables, la incertidumbre será máxima, y nos encontraremos en el límite superior de la entropía.**
- **Se mide en bits/símbolo. (***Obtenemos H bits por cada símbolo transmitido***).**

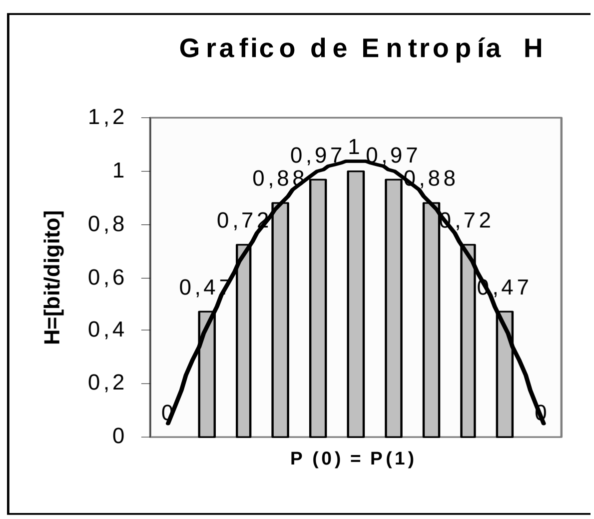
Entropía Normalizada

Para comparar la entropía de fuentes con diferente número de símbolos.

$$\hat{\mathrm{H}}(\mathrm{S}) = \frac{\mathrm{H(S)}}{\log_2(m)}$$ **de manera que** $0 \le \hat{\mathrm{H}}(\mathrm{S}) \le 1$.

Es como si estuviésemos normalizando en número de símbolos a 2 (número mínimo de símbolos).

$\mathrm{H_{max}} = \log m$ ->



Tasa de información (R)

Si dos fuentes tienen igual entropía, pero una es más rápida (produce más símbolos por unidad de tiempo): a igual tiempo, tranfiere más información la fuente más rápida.

Definición: Velocidad o Tasa de Información (R) es el Cociente entre la entropía de la fuente respecto de la duración promedio de los símbolos que ésta envía.

$$R = \frac{\mathrm{H}(s)}{\tau}$$ **bits por seg. (bps) ó Shannon por seg.**

τ = duración <u>promedio</u> de los símbolos (se mide en seg/símbolo).

R nos brinda la cantidad de información producida por una fuente, en un determinado tiempo.

Velocidad de señalización (r)

- **Las señales se transmiten en los canales como símbolos de duración fija.**
- **La duración de cada señal determina la Velocidad de Señalización.**
- **Es la inversa del ancho de pulso (ó símbolo) enviado:**

$$r = \frac{1}{\tau}$$ *[binits/seg.]=Baudio*

BAUDIO: el número de elementos de señal (símbolos) que puede ser transmitido por segundo.

También: la cantidad de dígitos o binits que es transmitida por segundo = *BAUDIOS*

*El **Baudio mide la velocidad de señalización**. Indica **<u>la velocidad de los símbolos</u>**.*

Unidad de la Tasa de Información

Usando la Velocidad de señalización, nuestra Tasa de información nos queda:

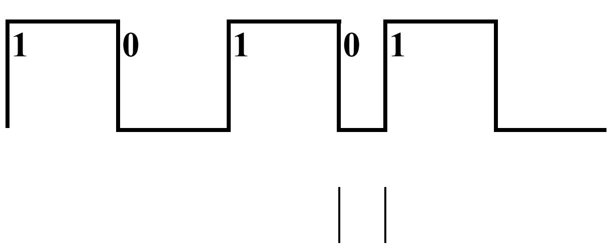
$$R = r \,.\, \mathrm{H}(s)$$

R = [binits/seg.] * [bits/binits] = [bits/seg.] (BPS)

Velocidad de señalización Máxima posible (S)

La inversa del periodo más corto de los pulsos que contenga el mensaje.

Con pulsos de señal de igual duración, r es el número de dichos pulsos por segundo, o el máximo número de transiciones de estados del canal por segundo.



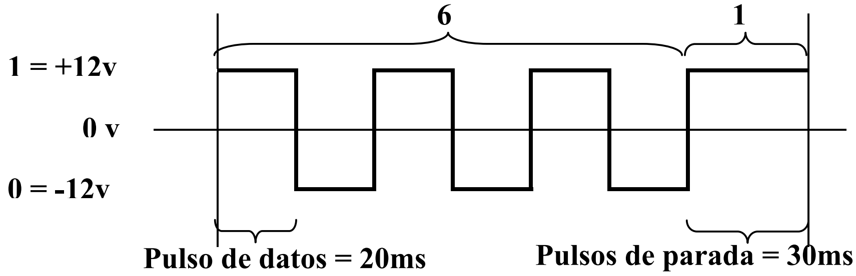
$T_{min.}$

$$S = \frac{1}{\mathrm{T}\min\,(seg)} \rightarrow S = (Baudios)$$

Es decir: r(max) = S

Ejemplo.

Emisión de datos codificados en pulsos bipolares con I = 1 Bit c/u (equiprobables**):**



$$T\min = 20ms$$

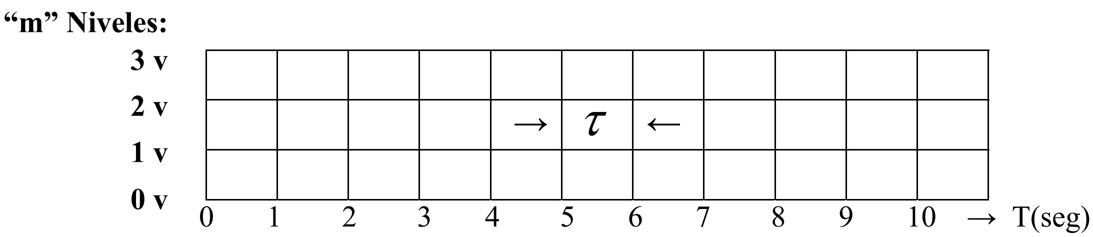
$$S = \frac{1}{20ms} = 50\,Baudios$$

Pero se transmiten siete binits cada 150 mseg.: $r = \dfrac{7\,binits}{150\ (mseg\ )} = 46\ binits\ /\ seg \Rightarrow 46\ Baudios$

Pero I_{stop}=0 => $R = \dfrac{6\,bits}{150\,(mseg)} = 40\,bits/seg$ **(Tasa o cantidad de información de la fuente)**

Si $R = r\ .\ \mathrm{H}(x)$ **->** $\mathrm{H} = \dfrac{R}{r} = \dfrac{40\ bits/seg}{46\ binits/seg} = 0{,}869\ bits/binits$ **(Inf. Prom. por c/símbolo entregado)**

CAPACIDAD DEL CANAL (C)



- **Posibilidades en cada intervalo: 4^1, 4^2, 4^3,…., $4^{(T/\tau)}$ → $m^{(T/\tau)}$**
- **La cantidad de símbolos distintos estará dada por la cantidad de niveles distintos.**

A Inf. Máx. (símbolos del mensajes equiprobables) la probabilidad de que aparezca un cierto símbolo en un tiempo "T**" será:** $\mathrm{P} = \dfrac{1}{m^{\mathrm{T}/\tau}}$ **y se transmite:** $\mathrm{I} = \log_x m^{\mathrm{T}/\tau} \rightarrow \mathrm{I} = \dfrac{\mathrm{T}}{\tau}\log_x m$

En un tiempo T, tranportará esa Información, por lo tanto (en bits):

$$\boxed{C = \frac{Información}{\mathrm{T}} \qquad \rightarrow \qquad C = \frac{1}{\tau}\log_2 m \qquad (bit/seg)}$$

Como $S = \dfrac{1}{\tau}$ **resulta:** $\mathrm{R = S\ log_x\ m}$

Lo cual significa que un canal que quiera transportar esa cantidad de información deberá tener una capacidad: C = S log m (o mayor):

$$\boxed{\mathbf{R \le C}}$$

CANAL IDEAL: Si intoducimos en la entrada *m* símbolos, a su salida reproduce exactamente los mismos símbolos. (En el canal ideal no se consideran ruido ni distorsión).

- **"<u>C = S log m" es la mayor información por unidad de tiempo que puede transportar</u>.**

TEOREMA DE NYQUIST

La velocidad de señalización esta relacionada con "B" (el *ancho de banda* **del canal).**

El teorema de Nyquist establece que:

$$S \leq 2\,B$$

Máxima velocidad de datos posibles en un canal sin ruido y con un ancho de banda infinito.

- ***Frecuencia de Nyquist*: Frecuencia de muestreo. No puede ser superior al doble de la frecuencia de la señal a muestrear: FN = 2 f (a mayor frecuencia, la inf. es "redundante").**

$$FN \;<=\; \mathbf{2\,\Delta F}$$

$$\textbf{Siendo } \Delta F = f_{(sup)} - f_{(inf)} = B$$

- **Canales sin ruido + Señales binarias + Transmisión mononivel: FN = máxima velocidad binaria. Es un límite físico, dado que se necesita oscilar para transmitir información.**
- **Transmisión multinivel: Cada símbolo contiene más bits y por lo tanto I >1 :**

$$C = \underbrace{2\,\mathbf{B}}_{\geq S}\ \underbrace{\log_2 m}_{H} \qquad \text{También:} \quad C = 2\,\mathbf{B.\ H}$$

<u>A esta velocidad binaria la llamamos Límite de Nyquist.</u>

TEOREMA DE SHANNON - HARTLEY

Hartley, y después Shannon, demostraron que la capacidad de un canal de transportar información en presencia de ruido vale:

$$\boxed{C = B \log_2 (1 + S/N) \quad [\text{bits/seg.}]}$$

B = ancho de banda del canal

S = potencia media de la señal

N = potencia media del ruido

S/N **es la relación señal a ruido expresada en <u>número de veces</u> (no en dB).**

Comparando con el *Teorema de Nyquist:* $C = 2\,\mathbf{B}\log_2 m \rightarrow C = \mathbf{B}\log_2 m^2$

Donde: $m^2 = 1 + S/N$

$$\boxed{m = \sqrt{1+\frac{S}{N}} = \frac{\sqrt{S+N}}{\sqrt{N}} = \frac{Valor\ \ eficaz\ de\ (la\ señal + Ruido)\,/\,1\Omega}{Valor\ eficaz\ del\ Ruido\,/\,1\Omega}}$$

- **Es decir, podemos determinar el número máximo de niveles de la señal o cantidad de símbolos que se pueden enviar, considerando el ruido presente en el canal.**
- **"C" depende de la relación señal a ruido (S/N) y del ancho de banda, y es independiente de la tasa de información de la fuente (su velocidad).**