Teoria da Informação - 060046 - UNISINOS

Aula 2 – 2/mar/2007

$$H = -\sum_{i=1}^{M} P_i \log_2 P_i \quad \text{(bits per symbol).}$$

Entropia de uma fonte de Informação (Shannon)

$H' = mH$ onde H= bits/símbolo H' = bits/seg m=símbolos/seg N símbolos

$$p = p_1^{p_1 N} p_2^{p_2 N} \cdots p_n^{p_n N}$$

$$\log p \doteq N \sum_i p_i \log p_i$$

$$\log p \doteq -NH$$

$$H \doteq \frac{\log 1/p}{N}.$$

$\frac{\log p^{-1}}{N}$ very close to H when N is large

Entropia relativa e redundância (Elwyn)

$$\text{Incerteza Relativa} = \frac{\text{Incerteza Real}}{\text{Incerteza Máxima (sendo } n = 4)}$$

$$= \frac{|\ \Sigma p_i \log p_i}{\log n}$$

$$= \frac{1\frac{3}{4}}{2} \text{ ou } \frac{7}{8} \text{ ou } 87{,}5 \text{ por cento.}$$

$$\text{Redundância} = \frac{\text{Incerteza Máxima} - \text{Incerteza Real}}{\text{Incerteza Máxima}}$$

$$= 1 - \text{Incerteza Relativa}$$

$$\text{Incerteza Relativa} = \frac{-\Sigma p_i \log p_i}{\log n}$$

$$= \frac{4{,}129}{4{,}700}$$

$$= 87{,}9 \text{ por cento}$$

Redundância $= 1 - 0{,}879$, ou seja, 12,1 por cento.

Freqüência e valor de informação de letras, no inglês escrito

Letra	p_i	$-\log p_i \times p_i$
A	0,082	0,296
B	0,014	0,086
C	0,028	0,144
D	0,038	0,179
E	0,131	0,384
F	0,029	0,148
G	0,020	0,113
H	0,053	0,225
I	0,063	0,251
J	0,001	0,010
K	0,004	0,032
L	0,034	0,166
M	0,025	0,133
N	0,071	0,271
O	0,080	0,292
P	0,020	0,113
Q	0,001	0,010
R	0,068	0,264
S	0,061	0,246
T	0,104	0,340
U	0,025	0,133
V	0,009	0,061
W	0,015	0,091
X	0,002	0,018
Y	0,020	0,113
Z	0,001	0,010

		PARA				Total de Fileira
		A	B	C	D	
DE	A	4	5	5	5	19
	B	9	1	0	0	10
	C	2	3	0	0	5
	D	5	0	0	0	5
Total da Coluna		20	9	5	5	

		Y				Total de Fileira
		A	B	C	D	
X	A	$\frac{4}{19}$	$\frac{5}{19}$	$\frac{5}{19}$	$\frac{5}{19}$	1
	B	$\frac{9}{10}$	$\frac{1}{10}$	0	0	1
	C	$\frac{2}{5}$	$\frac{3}{5}$	0	0	1
	D	1	0	0	0	1

$$-\sum_{Y=A}^{D} p_A(Y) \log p_A(Y), \text{ ou seja}$$

$$\left(-\frac{4}{19} \log \frac{4}{19}\right) + \left(-\frac{5}{19} \log \frac{5}{19}\right) + \left(\frac{5}{19} \log \frac{5}{19}\right) + \left(-\frac{5}{19} \log \frac{5}{19}\right)$$

X	$-\Sigma p_X(Y) \log p_X(Y)$	$p(X)$	Produto
A	1,994	0,5	0,997
B	0,469	0,25	0,117
C	0,971	0,125	0,121
D	0,000	0,125	0,000

$$H_{\mathbf{x}}(Y) = -\sum^{\mathbf{X}}\sum^{\mathbf{Y}} p(X)p_{\mathbf{x}}(Y)\ \log\ p_{\mathbf{x}}(Y) \qquad = 1.235 \text{ bits/letra}$$

$$\boxed{\begin{aligned} H(\text{digramas}) &= H(Y) - H_{\mathbf{x}}(Y) \\ &= 1{,}75 - 1{,}235 \ \text{bits/letra} \\ &= 0{,}415 \ \text{bits/letra.} \end{aligned}} \qquad \boxed{\begin{aligned} &\frac{1{,}235}{2} \times 100 \ \text{por cento} \\ &= 61{,}75 \ \text{por cento.} \end{aligned}}$$

redundância=1-0,6175=0,3825

<u>Representação de Operações de codificação e decodificação</u> (Shannon)

$$\boxed{\begin{aligned} y_n &= f(x_n, \alpha_n) \\ \alpha_{n+1} &= g(x_n, \alpha_n) \end{aligned}}$$

x_n is the n^{th} input symbol,

α_n is the state of the transducer when the n^{th} input symbol is introduced,

y_n is the output symbol (or sequence of output symbols) produced when x_n is introduced if the state is α_n.

<u>Canal de comunicação sem ruído (noiseless)</u>

Exemplo: todos símbolos com a mesma duração; 32 símbolos (5 bits); R= n símbolos/seg, então:

$\boxed{\text{C= 5n bits/seg onde C é a capacidade do canal}}$

$R_{max} = \dfrac{C}{H} - \epsilon$ onde R é a taxa de transmissão

<u>Equivocação e capacidade do canal</u>

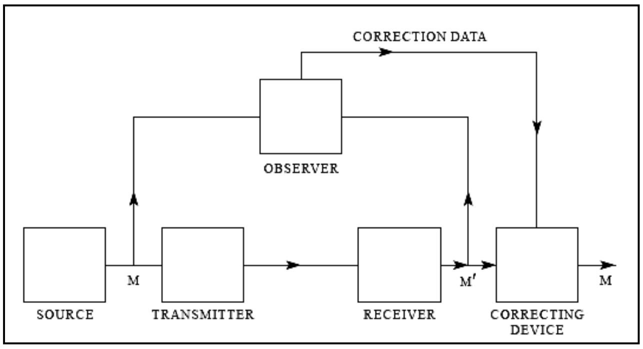
Exemplo: alfabeto={0,1} taxa de transm.= 1000 símb./seg $p_0=p_1= 1/2 = 0.5$ erro = 1%

$$R = H(x) - H_y(x)$$

p=0.99 e q=0.01 (p=1-q)

$$H_y(x) = -[.99\log.99 + 0.01\log 0.01]$$
$$= .081 \text{ bits/symbol}$$

R=1000-81=919 bits/seg



no caso extremo, se p=q= 1/2 = 0.5

$$H_y(x) = -\left[\tfrac{1}{2}\log\tfrac{1}{2} + \tfrac{1}{2}\log\tfrac{1}{2}\right]$$
$$= 1 \text{ bit per symbol}$$

e nesse caso R=1000 bits transmitidos – 1000 bits equivocados = 0

<u>Teorema fundamental do canal discreto com ruído</u>

$$\boxed{C = \text{Max}\big(H(x) - H_y(x)\big)}$$