Tema 2: Conceptes bàsics

Informació i Seguretat (Universitat Autònoma de Barcelona)

Adrian Tamayo Martinez

TEMA 2
Conceptes bàsics de la teoria de la informació

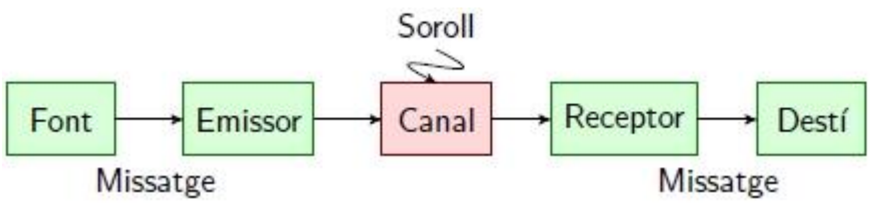
1. INTRODUCCIÓ

Com transmetre la informació o emmagatzemar-la:
- Forma **eficient**: minimitzant recursos (espai, temps).
- Forma **exacta**: sense pèrdua d'informació.
- Forma **segura**: sense manipulació de la informació

Solució => Teorema de la comunicació de C. E. Shannon (1948)

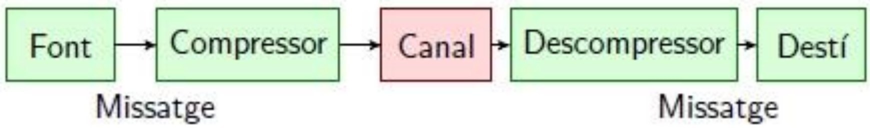
Teorema: teoria matemàtica que tracta de la transmissió, emmagatzematge i transformació de la informació.

Arquitectura d'un sistema de comunicació



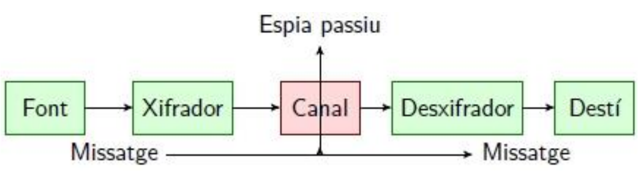
Font: persona / màquina que produeix informació.
Emisor / Codificador: adapta informació a les característiques del canal.
Canal: mitjà pel qual es transmet el missatge.
Soroll: error o pertorbació aleatòria característica del canal.
Receptor / Descodificador: recuperació de la informació (sovint amb errors).
Destí: persona o màquina que rep la informació.

Transmissió eficient

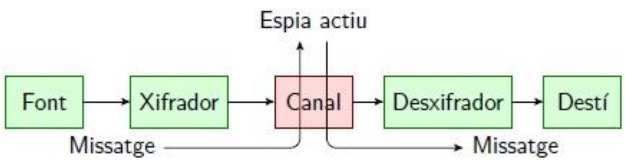


Espies: privacitat i autenticitat

Privacitat: preserva la confidencialitat de les dades (espia passiu).



Autenticitat: impedeix la modificació no autoritzada de les dades (espia actiu).



2. MESURA DE LA INFORMACIÓ

Incertesa: falta de certesa davant d'una determinada situació o experiment, que no s'ha realitzat anteriorment o amb resultats de caràcter aleatori.

Quantitat d'informació obtinguda de l'experiment = **quantitat d'incertesa** abans de l'experiment.

Exemple:
Quin resultat dóna més informació, treure un 5 a un dau perfecte o creu a una moneda perfecta?

- El resultat que dóna més informació és treure un 5 a un dau, ja que en el dau hi ha mes casos / esdeveniments (6 possibles) que a una moneda (2 possibles).

Quin resultat dóna més informació, treure cara amb p(cara) = 0,9 o creu amb p(creu) = 0,1 en una moneda perfecta?

- El resultat que dóna més informació es treure creu, ja que amb probabilitat p(cara) = 0,9 (90%) sabem que en un 90% dels cops hi sortirà cara, però només el 10% restant hi sortirà creu.

Funció d'incertesa

Diem I(n) a la incertesa sobre n resultats possibles i equiprobables.

Requisits:

- $I(1) = 0, i\ I(n) < I(n+1), \forall n \in N$
- $I(nm) = I(n) + I(m), \forall n, m \in N$
- $I(n^k) = k \cdot I(n), \forall n, k \in N$

En 1928, Hartley proposa => $I(n) = \log(n)$, que compleix els requisits anteriors.
Problema => no es tenen en compte les probabilitats de cada resultat.
Solució => Mesura de Shannon.

Propietats del logaritmes

1) $\log_x(1) = 0$
2) $\log_x(n \cdot m) = \log_x(n) + \log_x(m)$
3) $\log_x\left(\frac{n}{m}\right) = \log_x(n) - \log_x(m)$
4) $\log_x(a^n) = n \cdot \log_x(a)$
5) $\log_x\left(\frac{1}{n}\right) = -\log_x n$

3. MESURA DE SHANNON

En 1948, C. E. Shannon proposa => $I(A) = \log\left(\frac{1}{p(A)}\right) = -\log(p(A))$ com a mesura de la incertesa d'un esdeveniment A amb probabilitat p(A).

Informació d'una font

Si una font d'informació produeix símbols a1,a2,...,an amb probabilitats p(a1),p(a2),...,p(an), les informacions associades a aquests símbols dependran de la seva probabilitat.

La informació de la font serà la mitjana ponderada (esperança) de la informació de tots els símbols:

$$\sum_{i=1}^{n} p(a_i) \cdot I(a_i) = \sum_{i=1}^{n} p(a_i) \cdot \log\left(\frac{1}{p(a_i)}\right)$$

Unitats de mesura de la informació

Unitat d'informació més petita = bit. Quan volem la informació mesurada en bits, fem servir $\mathbf{\log_2(x)}$.

- Si la base és 10 => unitat s'anomena **dit** (o Hartley). Grau d'incertesa corresponent a 10 esdeveniments possibles i equiprobables.
- Si la base és el número e (cas continu), unitat de mesura => **nat.**

4. ENTROPIA D'UNA VARIABLE ALEATÒRIA DISCRETA

Sigui X una v.a. discreta amb distribució de probabilitats {p1,...,pn}, on $p_i > 0\ i\ \sum_{i=1}^{n} p_i = 1$.

Aleshores, l'entropia de X és: $H(X) = -\sum_{i=1}^{n} p_i \cdot \log(p_i) = \sum_{i=1}^{n} p_i \cdot \log\left(\frac{1}{p_i}\right)$

Exemple:

Sigui S = {a1,a2,a3} amb probabilitats p1 = 1/2, p2 = p3 = 1/4. Quant val l'entropia?

$$H(S) = \frac{1}{2} \cdot \log\left(\frac{1}{\frac{1}{2}}\right) + \frac{1}{4} \cdot \log\left(\frac{1}{\frac{1}{4}}\right) + \frac{1}{4} \cdot \log\left(\frac{1}{\frac{1}{4}}\right) = \frac{1}{2} \cdot \log(2) + \frac{1}{4} \cdot \log(4) + \frac{1}{4} \cdot \log(4)$$
$$= \frac{1}{2} \cdot \log(2) + \frac{2}{4} \cdot \log(4) = \frac{1}{2} + 1 = \mathbf{1,5}$$

Teorema fonamental de l'entropia

Sigui X una v.a. discreta amb distribució de probabilitats {p1,...,pn} aleshores:

- $H(X) \leq \log(n)$
- $H(X) = \log(n)\,, si\ i\ només\ si\ p_i = \frac{1}{n}, \forall i$

(*) L'entropia d'una v.a és màxima si la distribució de probabilitats és equiprobables.

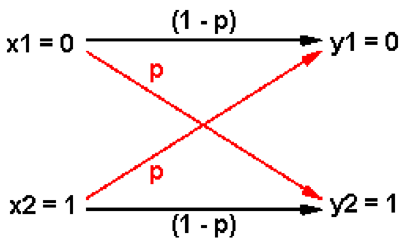
Entropia binària

Entropia d'una font que emet 0's i 1's.

$S = \{a1, a2\}, p(a1) = p, p(a2) = (1 - p)$

$$H(X) = H(p, 1 - p) = -p \cdot \log(p) - (1 - p).\log(1 - p)$$

Canal amb soroll



BSC => canal binari i simètric.

X = {x1 = 0, x2 = 1}
Y = {y1 = 0, y2 = 1}

p és la probabilitat d'error al bit.

Les probabilitats condicionades, p(y1|x1) = p(y2|x2) = 1 – p i p(y2|x1) = p(y1|x2) = p

Entropia conjunta de dues v.a. discretes

$$H(X,Y) = -\sum_{I=!}^{n}\sum_{j=1}^{m} p(x_i, y_j) \cdot \log(p(x_i, y_j)$$

Entropia condicionada d'X donat Y = y

$$H(X, Y = y_i) = -\sum_{i=1}^{n} p(x_i|y_j) \cdot \log(p(x_i|y_j)$$

$$H(Y, X = x_i) = -\sum_{j=1}^{n} p(y_j|x_i) \cdot \log(p(y_j|x_i)$$

Entropia condicionada d'X donat Y

La incertesa que tenim respecte a l'entrada sabent la sortida serà (entropia condicionada d'X, donat Y):

$$H(X,Y=y_i) = \sum_{j=1}^{m} p(y_j) \cdot H(X|Y=y_i) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i|y_j) \cdot \log(p(x_i|y_j)$$

La incertesa respecte a la sortida sabent l'entrada és (entropia condicionada d'Y donat X):

$$H(Y,X) = \sum_{i=1}^{n} p(x_i) \cdot H(Y,X=x_i) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(y_j|x_i) \cdot \log(p(y_j|x_i)$$

Propietats entropia condicionada

- H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y).
- H(X,Y) < o = H(X) + H(Y) => només sí X i Y són independents.
- H(X|Y) < o = H(X), H(Y|X) < o = H(Y) => només sí X i Y són independents.
- H(X) – H(X|Y) = H(Y) – H(Y|X).

Informació mútua entre dues v.a. discretes

I(X,Y) = H(X) – H(X|Y) = H(Y) – H(Y|X)

Capacitat d'un canal

$$C = \max\big(I(p1, \dots, pn)\big)$$

- La capacitat indica la quantitat màxima d'informació que pot passar per símbol d'entrada, mesurat en bits/símbol o bits/entrada.