NETWORK INFORMATION THEORY

Studia i limiti fondamentali della rappresentazione e della comunicazione di dati in una rete

Teoria classica: una sorgente
una destinazione

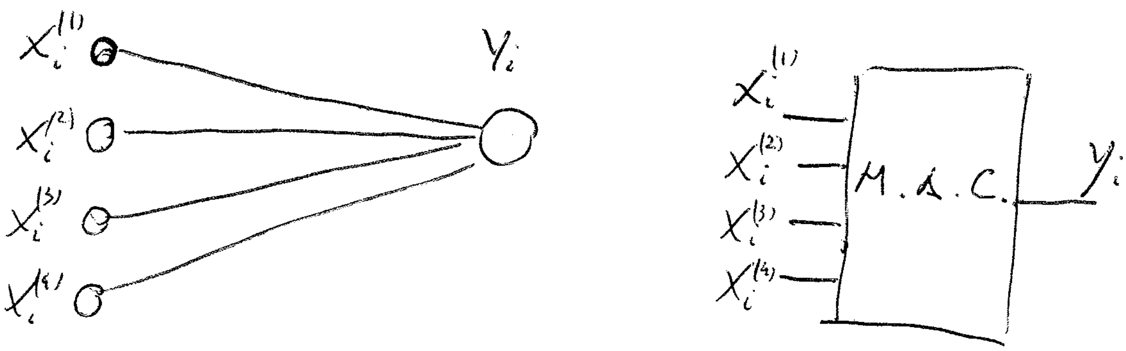
NIT: una/molte sorgenti
una/molte destinazioni

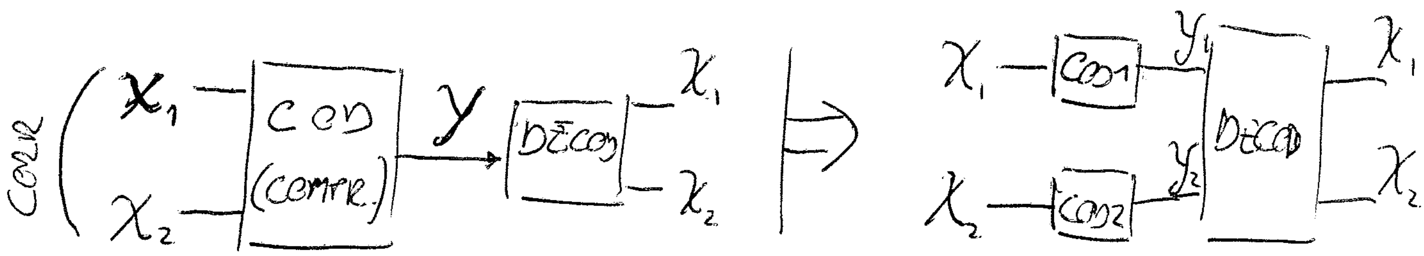
Applicazioni: la rappresentazione e la trasmissione di dati è fondamentale in prat. tutte le applicazioni

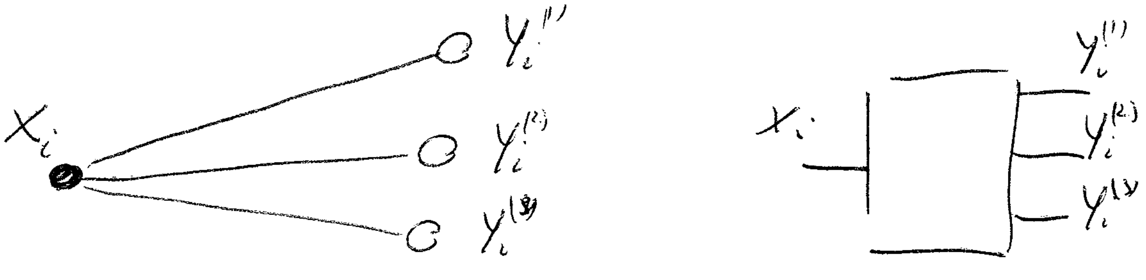
- telecom:
  - reti di sensori
  - reti radiomobili
- elettronica:
  - memorizzazione efficiente
  - modellizzazione reti intra e inter chip efficienti
- CS:
  - IPC efficiente
  - Reti scalabili
  - P^2P

Overview

- Teoria classica singola sorgente singola destinazione
- - STRUMENTI e concetti
- canale ad accesso multiplo
- codifica di sorgente distribuita
- canale "broadcast"
- RETI GENERICHE: new Trends and analysis Techniques







Entropia Informazione Mutua ... Proprietà

$X \quad \mathcal{X} = \{x_1, \dots, x_m\} \quad p = (p_1, \dots, p_m)$

- $H(X) = -\sum p_i \log p_i = E\{\log p(X)\}$ ENTROPIA
- $H(X) \geq 0 \qquad H(X) \leq \log m$ (R.V. discrete)

$(X, Y) \quad \mathcal{X} \quad \mathcal{Y}$

- $H(X,Y) = -\sum_{x_i}\sum_{y_i} p(x_i, y_i) \log p(x_i, y_i) = E\{-\log p(X,Y)\}$

ENTROPIA CONGIUNTA

- $H(Y|X) = H(X,Y) - H(X)$ ENTROPIA CONDIZIONATA
- $I(X;Y) = H(X) - H(X|Y) =$
  $= H(Y) - H(Y|X)$

INFORMAZIONE MUTUA FRA X e Y

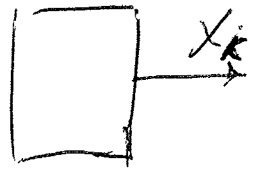
- $I(X;Y) = I(Y;X) \geq 0$ (= se indip)
- $H(X) \geq H(X|Y)$ (= se indip)
- $D(p\|q) = \underbrace{\sum p_i \log \frac{p_i}{q_i}}_{E_p\{\log \frac{p_i}{q_i}\}} \geq 0$ (=0 se ≡)

ENTROPIA RELATIVA

- $I(X;Y) = D\big(p(x,y) \| p(x)p(y)\big)$

Teoria dell'Informazione : Risultati classici

Strumento : Asymptotic Equipartition Property



Sorgente

X_k R.V.

$\mathcal{X} = \{x_1, \dots, x_m\}$

p.m.f. $p(x_i) = p_i = Pr\{X_k = x_i\}$

$\{X_k\}$ processo stocastico

Come sono le sequenze in uscita dalla sorgente?

Come si distribuisce la probabilità delle sequenze?

Osservo $X_1, \dots, X_n$ suppongo i.i.d.

$$\underbrace{p(X_1, \dots, X_n)}_{R.V.} = \prod_{i=1}^{n} \underbrace{p(X_i)}_{R.V.}$$

$Y_i = p(X_i)$ hanno tutte la stessa p.m.f. (i.i.d.)

$$-\frac{1}{n} \log p(X_1, \dots, X_n) = \frac{1}{n} \sum \underbrace{-\log p(X_i)}_{R.V.\ (i.i.d.)}$$

$n \to \infty$ $\quad \frac{1}{n} \sum -\log p(X_i) \to \boxed{E\{\log p(X_i)\}}$ $\quad Pr = 1$

L.G.N.

è un numero che caratterizza $p(X_1, \dots, X_n)$

$$H(X) = E\{-\log_2 p(X)\}$$ ENTROPIA

SEQUENZE TIPICHE

$$2^{-n(H(X)+\varepsilon)} \leq p(X_1, \dots, X_n) \leq 2^{-n(H(X)-\varepsilon)}$$

con elevata probabilità (vicina a 1 a piacere al crescere di n)

<u>Insieme Tipico</u> (o delle sequenze tipiche di lunghezza n con parametro ε)

$$A_\varepsilon^{(n)} = \left\{ (x_1, \dots, x_n) : 2^{-n(H(X)+\varepsilon)} \leq p(x_1, \dots, x_n) \leq 2^{-n(H(X)-\varepsilon)} \right\}$$

È un concetto potente che racchiude gran parte dei concetti della IT

$A_\varepsilon^{(n)}$ ha circa $2^{nH(X)}$ elementi $\left(|A_\varepsilon^{(n)}| \simeq 2^{nH(X)}\right)$

$$\underbrace{(1-\varepsilon)2^{n(H(X)-\varepsilon)}}_{n > n^*} \leq |A_\varepsilon^{(n)}| \leq 2^{n(H(X)+\varepsilon)}$$

Applicazione: Compressione dei dati

$X = \{x_1, \dots, x_m\}$

n° di sequenze di lunghezza n : m^n

n° di sequenze tipiche : $2^{nH(X)}$

Rappresentazione banale $n \log_2 m$ bit $\quad H(X) \leq m$

" efficiente $nH(X)$ bit

- $P_r\{A_\varepsilon^{(n)}\} > 1-\varepsilon$ $\quad n$ suff grande

L.G.N. :

$\forall \varepsilon > 0 \quad \exists n^* / \quad \forall n > n^*$

$$P_r\left\{ \left| \frac{1}{n}\sum Y_i - EY \right| < \varepsilon \right\} > 1-\varepsilon$$

- $|A_\varepsilon^{(n)}| \leq 2^{n(H(X)+\varepsilon)}$

$$1 = \sum_{\underline{x}\in\mathcal{X}^n} p(\underline{x})$$

$$\geq \sum_{\underline{x}\in A_\varepsilon^{(n)}} p(\underline{x})$$

$$\geq \sum_{\underline{x}\in A_\varepsilon^{(n)}} 2^{-n(H(X)+\varepsilon)}$$

$$= 2^{-n(H(X)+\varepsilon)} |A_\varepsilon^{(n)}|$$

- $|A_\varepsilon^{(n)}| \geq (1-\varepsilon)\, 2^{n(H(X)-\varepsilon)}$ $\quad$ n suff. grande

$$1-\varepsilon < P_r\{A_\varepsilon^{(n)}\}$$

$$= \sum_{\underline{x}\in A_\varepsilon^{(n)}} p(\underline{x})$$

$$\leq \sum_{\underline{x}\in A_\varepsilon^{(n)}} 2^{-n(H(X)-\varepsilon)}$$

$$= 2^{-n(H(X)-\varepsilon)} |A_\varepsilon^{(n)}|$$

RISULTATI IMPORTANTI: COMPRESSIONE DEI DATI

$H(X)$ caratterizza una RV o un processo i.i.d.

$H(\mathcal{X}) = \lim_{n\to\infty} \frac{1}{n} H(X_1,\dots,X_n)$ ENTROPY RATE

Se esiste caratterizza un processo stocastico

Codifica di Sorgente:

$x_1,\dots,x_n \longrightarrow C^{(n)}(x_1,\dots,x_n)$ sequenza binaria

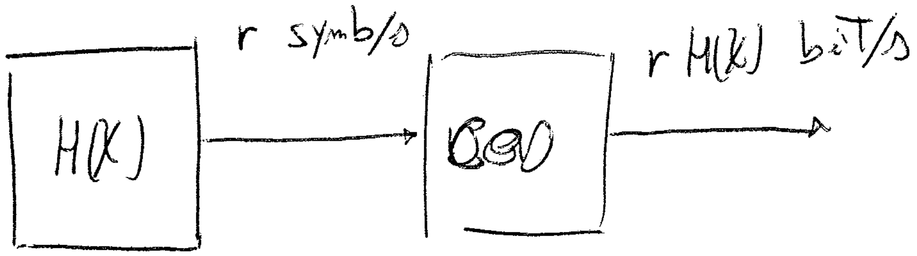
$L(C)$ numero medio di bit per simbolo X_i

Th:

$$L(C^{(n)}) \geq H(\mathcal{X})$$

Inoltre, $\forall \varepsilon > 0 \quad \exists\, n^* \text{ e } \tilde{C}^{(n^*)} \;|$

$$L(\tilde{C}^{(n^*)}) \leq H(\mathcal{X}) + \varepsilon$$



TRASMISSIONE AFFIDABILE

Canale senza memoria (DMC)

$$p(y_k|x_k, y_1,\dots,y_{k-1}) = p(y_k|x_k) \quad \text{(con feedback)}$$

e anche

$$p(y_1,\dots,y_n|x_1,\dots,x_n) = \prod_{i=1}^{n} p(y_i|x_i)$$

- Il ~~massimo~~ supremo dei tassi R di informazione che si può trasferire in modo affidabile dal in all'out è

$$I(X;Y)$$

con una data distribuzione statistica dell'ingresso.

- Ottimizzando la distribuzione di ingresso $p(x)$ si ha

$$C = \max_{p(x)} I(X;Y)$$ capacità di canale

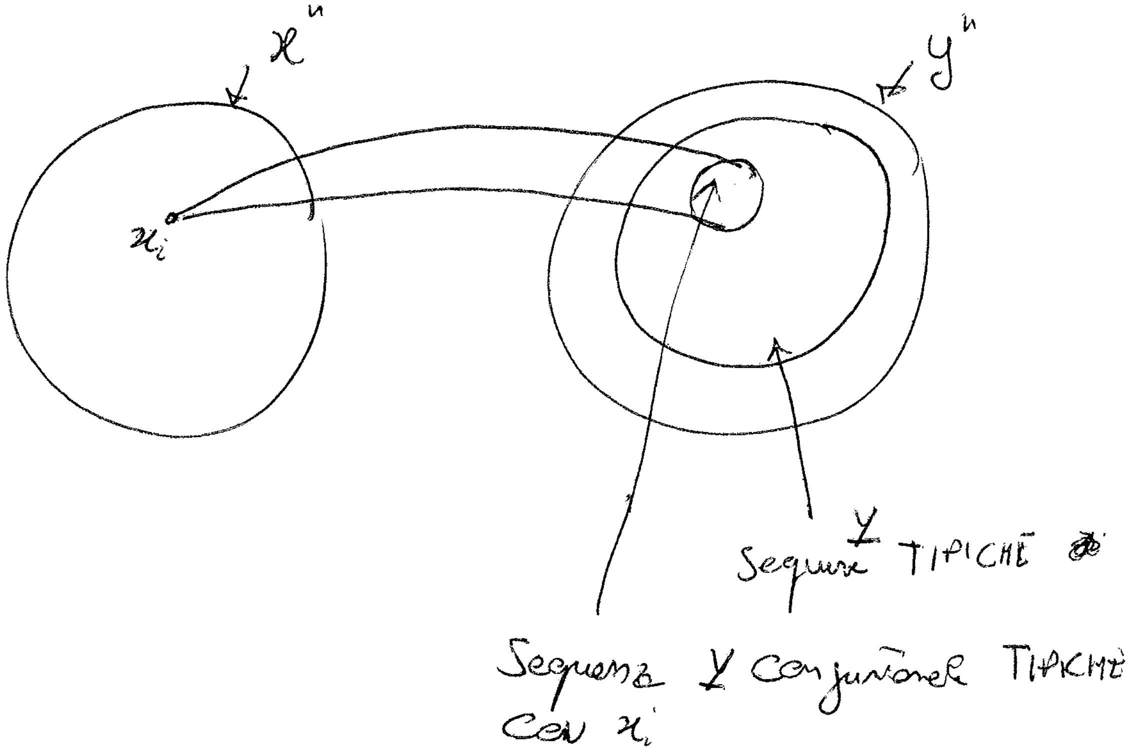
La TX affidabile si ottiene con un <u>CODICE</u>
che associa un indice da TX w
$w \in (1,\dots,2^{nR})$ ad un vettore di simboli ~~da~~
$x(w) = (X_1,\dots,X_n)$ da TX

R è il TASSO DEL CODICE

Un tasso R si dice RAGGIUNGIBILE se esiste una sequenza di leggi di codifica di tasso R $C(n)$ che garantisce una probabilità di errore piccola a piacere su qualunque parola di codice al crescere di n

Th: • $\forall R < C$ è RAGGIUNGIBILE

• $\forall R > C$ NON è RAGGIUNGIBILE



Sequenze Tipiche $\underline{Y} \simeq 2^{nH(Y)}$

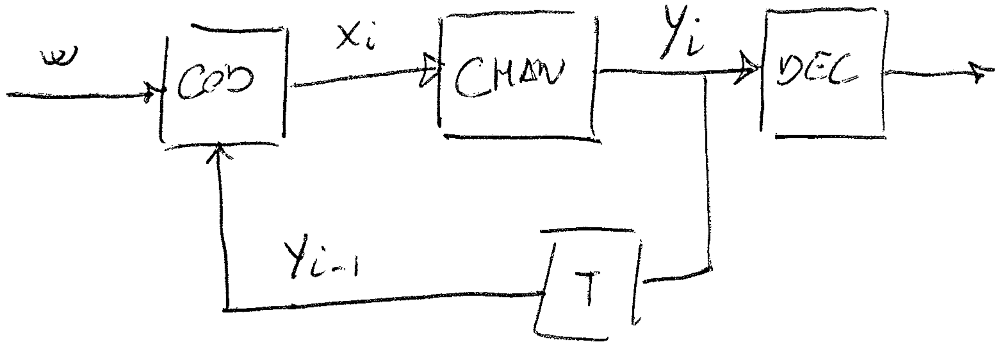
Sequenze Tipiche $\underline{Y}$ congiunte a $\underline{X} \simeq 2^{nH(Y|X)}$

max numero di sequenze x che danno origine a insiemi disgiunti

$$2^{nR} \leq \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{nI(X;Y)}$$

$\rightarrow R \leq I(X;Y)$

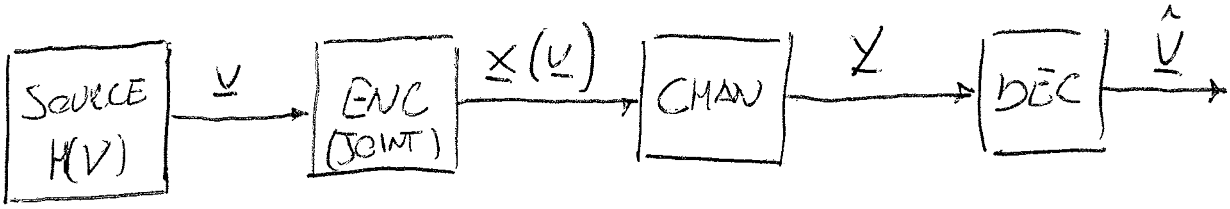
CAPACITÀ CON FEEDBACK



Th:

$$C_{FB} = C = \max_{p(x)} I(X;Y)$$

SEPARABILITÀ FRA CODIFICA DI SORGENTE E CODIFICA DI CANALE

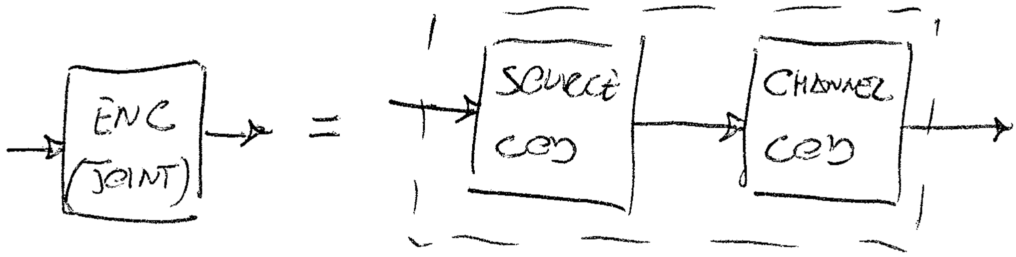


~~LA TRASMISSIONE PUÒ ESSERE RESA AFFIDABILE~~

~~SE E SOLO~~ SE

$$H(V) < C$$

LA TRASM. PUÒ ESSERE RESA AFFIDABILE ED È SUFFICIENTE CHE



VICEVERSA, SE

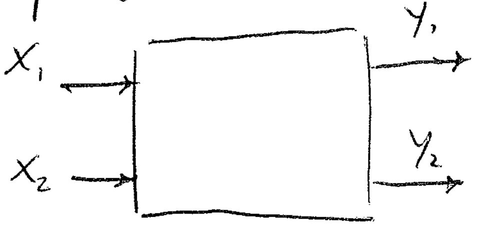
$$H(V) > C$$

LA TRASM. NON PUÒ ESSERE RESA AFFIDABILE

A.E.P. e SEQUENZE CONGIUNTAMENTE TIPICHE

Avremo a che fare con n-uple di simboli

ad esempio (canale con interferenza)



$$(X_1^{(1)} X_2^{(1)} Y_1^{(1)} Y_2^{(1)}) \; (X_1^{(2)} X_2^{(2)} Y_1^{(2)} Y_2^{(2)}) \dots$$

Tempo $\longrightarrow$

INDICE $\downarrow$

$$\begin{array}{c} X_1 \\ \vdots \\ X_k \end{array} \quad \begin{pmatrix} X_1^{(1)} \\ \vdots \\ X_k^{(1)} \end{pmatrix} \cdots \begin{pmatrix} X_1^{(n)} \\ \vdots \\ X_k^{(n)} \end{pmatrix}$$

Sia S un sottoinsieme di $\{X_1, \dots, X_k\}$ $\quad S \subseteq \{X_1, \dots, X_k\}$

$S_1, \dots, S_n$ realizzazioni i.i.d.

$$p(S_1, \dots, S_n) = \prod_{i=1}^{n} p(S_i)$$

$$-\frac{1}{n} \log p(S_1, \dots, S_n) \longrightarrow H(S)$$

L'insieme $A_\varepsilon^{(n)}$ di tutte le sequenze ε-tipiche di lunghezza n di vettori $(\underline{x}_1, \ldots, \underline{x}_k)$ è dato dal seguente

$$A_\varepsilon^{(n)}(X_1, \ldots, X_k) = \left\{ (\underline{x}_1, \ldots, \underline{x}_k) : \left| -\frac{1}{n} \log p(\underline{s}) - H(S) \right| < \varepsilon, \ \forall S \subseteq \{X_1, \ldots, X_k\} \right\}$$

è il vettore sottoinsieme di $(x_1, \ldots, x_k)$ secondo S

ad esempio se $S = \{ X_1, X_3, X_5 \}$

$$\underline{s} = \begin{pmatrix} x_1^{(1)} \\ x_3^{(1)} \\ x_5^{(1)} \end{pmatrix} \cdots \begin{pmatrix} x_1^{(n)} \\ x_3^{(n)} \\ x_5^{(n)} \end{pmatrix}$$

$$\underline{x}_i = x_i^{(1)} \ldots x_i^{(n)}$$

Th: $\forall \varepsilon > 0 \quad \exists n^* \ / \ \forall n > n^*$

- $P(A_\varepsilon^{(n)}) \geq 1 - \varepsilon$
- $\underline{s} \in A_\varepsilon^{(n)}(S) \rightarrow 2^{-n(H(S)+\varepsilon)} \leq p(\underline{s}) \leq 2^{-n(H(S)-\varepsilon)}$
- $S_1, S_2 \subseteq \{ X_1, \ldots, X_n \}$ se $(\underline{s}_1, \underline{s}_2) \in A_\varepsilon^{(n)}(S_1, S_2)$

$$2^{-n(H(S_1/S_2)+2\varepsilon)} \leq p(\underline{s}_1/\underline{s}_2) \leq 2^{-n(H(S_1/S_2)-2\varepsilon)}$$

Th: $S_1, S_2 \subseteq \{X_1, \dots, X_k\}$

$A_\varepsilon^{(n)}(S_1 | \underline{s}_2)$ è l'insieme delle sequenze $\underline{s}_1$ tipiche cong. con la sequenza $\underline{s}_2$. Se $\underline{s}_2 \in A_\varepsilon^{(n)}(S_2)$ per n sufficientemente grande

$$\left| A_\varepsilon^{(n)}(S_1 | \underline{s}_2) \right| \leq 2^{n\left(H(S_1|S_2) + 2\varepsilon\right)}$$

inoltre

$$(1-\varepsilon)\, 2^{n\left(H(S_1/S_2) - 2\varepsilon\right)} \leq \sum_{\underline{s}_2} p(\underline{s}_2) \left| A_\varepsilon^{(n)}(S_1 | \underline{s}_2) \right|$$

Th: $A_\varepsilon^{(n)}$ associato alla pmf $p(s_1, s_2, s_3)$

se si considera una sequenza $\tilde{S}$ associata a

$$\tilde{p}(s_1, s_2, s_3) = p(s_1|s_3)\, p(s_2|s_3)\, p(s_3)$$

$$P\{\tilde{\underline{S}} \in A_\varepsilon^{(n)}\} \leq 2^{-n\left(I(S_1; S_2 | S_3) - 6\varepsilon\right)}$$