Incertezza di un esperimento finito

Def: consideriamo un exp. X con K possibili ris. $\{x_1, \dots, x_K\}$ con probabilità $p(x_1), \dots, p(x_K)$, l'entropia è una misura dell'incertezza associata all'exp.

I 3 assiomi dell'entropia:

① $H(p_1, \dots, p_K)$ è una funz. continua delle p_i.

② Dati A e B con rispettiv. K e $K+1$ possibili ris. equiprobabili

$$\boxed{H(\underbrace{\tfrac{1}{K}, \dots, \tfrac{1}{K}}_{K \text{ volte}}) < H(\underbrace{\tfrac{1}{K+1}, \dots, \tfrac{1}{K+1}}_{K+1 \text{ volte}})}$$

(B è + incerto di A)

③ Dato X scomposto in due fasi:
1 — Y con $h < n$ ris. y_j, $j = 1, \dots h$.
2 — uno tra h exp. $Z(j)$, $j = 1, \dots h$ con n_j possibili ris. $(n_1 + \dots + n_h = n)$

$$\boxed{H(X) = H(Y) + \sum_{j=1}^{h} p(y_j) H(Z(j))}$$

Lemma 1.1

Se f è positiva e monotona crescente per cui vale che:
$\forall\, t > 1$ e $\forall\, n$ intero vale $f(t^n) = n f(t)$ allora f è della forma:

$$\boxed{f(t) = c \log(t)}$$

Lemma 1.2

$$\boxed{H(Y^n) = n\, H(Y)}$$

Y ripetuto n volte

Lemma 1.3

L'incertezza associata ad un exp. con K uscite equiprob. vale:

$$\boxed{c \log K}$$

Teorema

L'unica funzione H che soddisfa i 3 assiomi è:

$$\boxed{H(p_1, \dots, p_k) = -c \cdot \sum_{i=1}^{k} p_i \cdot \log p_i}$$

Def: L'incertezza associata ad un exp. X è detta entropia di X e vale:

$$\boxed{H(X) = -c \sum_{i=1}^{k} p(x_i) \cdot \log p(x_i)}$$

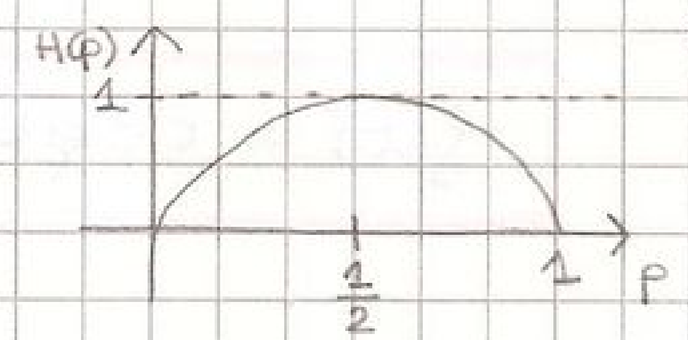
assumiamo: $\boxed{c = \frac{1}{\log 2}}$

Def: Un Bit è la quantità di incertezza associata ad un exp. con 2 sole uscite equiprobabili: $\boxed{H(1/2, 1/2) = 1}$

OSS:
- Se uno dei p_i vale $\emptyset$ allora la relativa componente della sommatoria vale $\emptyset$.
- Se uno dei p_i vale 1 allora tutti gli altri valgono $\emptyset$: $\sum_{i=1}^{k} p_i = 1$ $\Rightarrow$ l'info è nulla.
- Nel caso di 2 sole possibili uscite l'entropia vale:

$$\boxed{p \log p + (1-p) \log (1-p)}$$

Il grafico di $H(p, 1-p)$:



Lemma del logaritmo

Date 2 distrib. di prob.: $(p_1, \dots, p_k)$ e $(q_1, \dots, q_k)$ vale che:

$$\boxed{-\sum_{i=1}^{k} p_i \log p_i \le -\sum_{i=1}^{k} p_i \log q_i} \quad \text{con } \sum p_i = 1 \text{ e } \sum q_i = 1$$

il minimo si ha quando i p_i sono tutti uguali ai q_i

Proprietà dell'entropia

Teorema 1.2

Sia X un exp. con K possib. ris. e distribuzione di prob. $\{p_1, \dots, p_K\}$

vale che:

$$\boxed{0 \leq \underset{H(p_1,\dots,p_K)}{\underset{\shortparallel}{H(X)}} \leq \log K}$$

- $H(p_1, \dots, p_K) = 0$ se uno dei $p_i = 0$
- $H(p_1, \dots, p_K) = \log K$ se tutti i p_i sono uguali a $\frac{1}{K}$

Teorema

Siano X e Y due exp con rispettiv. K e h uscite.

Se con (X,Y) si indica il loro exp. congiunto, vale che:

$$\boxed{H(X,Y) \leq H(X) + H(Y)}$$

dove l'uguaglianza vale se X e Y sono INDIPENDENTI

Entropia condizionata

Def: Dati 2 exp. X e Y, l'incertezza associata all'exp. X una volta noto il ris. di Y è detta entropia condizionata di X dato Y e vale che:

$$\boxed{H(X|Y) = \sum_{s=1}^{h} p(y_s) \cdot H(X|y_s)}$$

Teorema

Dati due exp. X e Y vale che:

a) $H(X,Y) = H(Y) + H(X|Y)$

b) $H(X,Y) = H(X) + H(Y|X)$

c) $H(X|Y) \leq H(X)$ - - - - - - - -> = vale se X e Y sono INDIPENDENTI

Informazione reciproca

Def: Dati due exp X e Y la quantità non negativa:

$$\boxed{I(X|Y) = H(X) - H(X|Y)}$$

→ riduzione dell'incertezza nel ris. di X data la conoscenza del ris. di Y

è detta informazione reciproca tra i due exp.

OSS: $I(X|Y) \geq 0$ con = nel solo se i 2 eventi sono INDIPENDENTI

Sequenze tipiche ed equipartizione asintotica

Def: Sequenze di risultati che soddisfano una distribuzione proporzionale alle loro probabilità sono dette tipiche

Teorema della ripartizione asintotica

Dato un exp. X con $\{x_1, \dots x_k\}$ possibili risultati, e con prob. di uscita $\{p_1, \dots, p_k\}$ ed entropia $H(X)$, le sequenze di uscita di n esperimenti ripetuti $(n \to \infty)$, si possono dividere in 2 classi:

1) circa $2^{nH(X)}$ sequenze tipiche con prob. tà uniforme: $2^{-nH(X)}$
2) tutte le altre

Ma poiché la prob. di ottenere una seq. tipica è ~ 1, la prob. di ottenere una seq. NON tipica tende a 0.

SORGENTI DISCRETE

Classificazione delle sorgenti:

Una sorgente rappr. un modello di come vengono generati i dati che vogliamo trasmettere.

Def: Una **sorgente S** su un alfabeto $A = \{a_1, \dots, a_r\}$ è un dispositivo che ad ogni istante "t" emette un simbolo $X_t \in A$.

OSS: ogni poss. emissione può essere vista come un esp. X_i i cui poss. esiti appartengono ad A, le emissioni possono essere viste come una seq. di esp. casuali, MA non è lo stesso esp. ripetuto.
(se l'esp. è sempre lo stesso è un caso particolare)

Def: Una sorgente è **stazionaria** se $\forall$ simbolo generato e $\forall n$ la prob. che una sequenza $X_{t+1}, \dots, X_{t+n}$ appartenga alla stringa di uscita è indipendente da "t".
→ ciò ci permette di astrarre dal fattore tempo.

Def: Una sorgente è **senza memoria** se $\forall$ istante t e $\forall$ simbolo "a_i" generato da S, la prob.tà che X_t restituisca "a_i" NON dipende dai simboli $X_0, \dots, X_{t-1}$ usciti in precedenza.

Def: Una sorgente ha **memoria finita** se $\exists$ un intero m t.c. la prob.tà del simbolo in uscita dipende al più dagli m simboli usciti in precedenza.
Il minimo intero m per cui tale prop.tà è verificata è detto **memoria** di S.

Def: Data una sorg. stazionaria S con mem. "m" e prob.tà $\{p_1, \dots, p_r\}$ di generare $\{a_1, \dots, a_r\}$ la **sorgente adiacente $\bar{S}$** è una sorgente senza memoria con le stesse probabilità.

Def: Data una sorgente S, la sorgente **estensione k-esima** di S si indica con S^k e si ottiene considerando la sorgente con r^k simboli ottenuta raggruppando k a k i simboli emessi

OSS: La sorg. adiacente è un'altra sorg. rispetto ad S, a meno che S sia già senza memoria, l'estens. k-esima è la stessa sorgente in cui consideriamo le uscite raggruppate.

Teorema:

Se S ha memoria m, S^K ha memoria $\lceil \frac{m}{K} \rceil$

OSS: m non si azzera mai!!!

Def: Data una sorgente S stazionaria (con mem. "m" o ∞), la sorgente approssimazione di S di ordine K si indica con S_K e si ottiene considerando S con al più mem $= K$ e che ha le stesse prob.tà di S di emettere un simbolo dati i K precedenti.

Sorgenti con memoria finita e catene di Markov

AUTOMI:

- ASF: automi in cui gli insiemi dei possibili valori di ingresso, uscita e stato sono insiemi finiti.
- ASF det: $\forall$ coppia <stato, ingresso> c'è una sola possibile uscita
- ASF non det: $\forall$ coppia <stato, ingresso> ci possono essere più uscite e devo percorrere tutti i poss. cammini. Basta che uno di questi ci porti in uno stato finale per avere successo.
- ASF-D e ASF-ND sono equivalenti ci sono modi per passare da D a ND.

(MIA)

Def: Una catena di Markov è un automa probabilistico in cui per ogni stato possiamo avere più scelte percorse in base ad una certa prob.tà.
ogni stato ha una certa prob.tà di essere raggiunto (anche nulla).

In pratica:

Data una sorg. S con $A = \{a_1, \dots, a_r\}$ e mem. m, il cui comportamento dipende dagli ultimi m simboli, nelle catene di Markov si associa ad ogni possibile m-upla uno stato.

$\Rightarrow$ • $\sigma_1, \dots, \sigma_s$ possibili stati con $s = r^m$

• $p(a_j | \sigma_i)$ con $j = 1, \dots, r$ e $i = 1, \dots, r^m$ (sono r^{m+1})

↳ prob.tà che esca a_j se sono nello stato σ_i

Date queste prob.tà condizionate è possibile costruire la catena di Markov, infatti quando emetto un nuovo simbolo passo in un nuovo stato.

⇒ Alla catena di Markov è associata la matrice π di elementi π_{ij} e di dimensione $r^m \times r^m$ in cui ogni elemento è di tipo:

$$\pi_{ij} = p(b_j | b_i)$$

↳ prob.tà di passare dallo stato b_i allo stato b_j

OSS: la dim è $r^m \times r^m$, ma per ogni stato posso avere solo r possibili ingressi, quindi ci saranno molti "0" (stati in cui non posso andare)

⇒ $\forall$ riga al più r caselle $\neq$ "0".

⇒ la somma delle prob.tà delle righe è = "1".

⇒ la matrice di Markov è detta **stocastica**, poiché vale sempre:

1) $\pi_{ij} \geq 0$

2) $\sum_j \pi_{ij} = 1$

Def: la prob.tà di trovarmi, al tempo "t", nello stato "b_j" vale:

$$\boxed{p(b_j^{(t)}) = \sum_{i=1}^{n} p(b_i^{(t-1)}) \cdot p(b_j | b_i) = \sum_{i=1}^{n} p(b_i^{(t-1)})\, \pi_{ij}}$$

Def: la prob.tà di avere l'uscita "a_j" al tempo "t" vale:

$$\boxed{p(a_j^{(t)}) = \sum_{i=1}^{n} p(b_i^{(t-1)}) \cdot p(a_j | b_i)}$$

Nota:

La catena di Markov che schematizza la sorgente può essere definita da un **grafo associato alla matrice π**.

Distribuzione stazionaria di una sorgente di Markov

Def: Con $p(s_i^{(0)})$ indichiamo la prob.tà di trovarsi nello stato s_i all'istante "0", ossia quando la sorgente parte.

Def: Indichiamo con X_0 il vettore delle probabilità degli stati all'istante "0"

⟹ All'istante successivo la distribuzione di probabilità degli stati sarà data dal vettore X_1^T ottenuto moltiplicando X_0^T per Π:

trasposto

$$X_1^T = X_0^T \Pi$$

(MIA)
OSS: $X_0^T = \left(p(s_1^{(0)}), \dots, p(s_n^{(0)}) \right)$

potenza t-esima della matrice di Markov

⟹ In generale: $X_t^T = X_{t-1}^T \cdot \Pi = X_0^T \cdot \Pi^t$

Def: Indichiamo con X_t^T il vettore delle probabilità degli stati all'istante "t" in una catena di Markov.

AUTOVALORI E AUTOVETTORI NELLE MAT. STOCASTICHE:

- ci sarà sempre un autovalore $\lambda = 1$
- l'autovettore relativo avrà tutti gli elementi uguali $\begin{pmatrix} a \\ \vdots \\ a \end{pmatrix}$
- $\rho = 1$ poiché:
  - $|\lambda_i| \leq \rho \leq \|\Pi\|_\infty$
  - il massimo dei $|\lambda_i|$ è 1
  - $\|\Pi\|_\infty = 1$, poiché $\sum_j \Pi_{ij} = 1$

  RO (raggio spettrale $= \max |\lambda_i|$)

  Non ci sono autovalori più grandi di 1
- Π^T ha gli stessi autovalori e lo stesso ρ, quindi $\exists$ un autovettore per cui vale: $\Pi^T X = X$

  ⟹ $\exists\ w^T \mid w^T = w^T \Pi$

Teorema 2.2

La matrice di Markov Π ha almeno un autovalore "1", tutti gli altri autovalori sono in modulo NON maggiori di "1".
(Ovvero il raggio spettrale $\rho(\Pi) = 1$).

Supp. che S abbia distrib. di probabilità W all'istante "0", se vale che: $W^T \Pi = W^T$ allora la distrib. rimane la stessa all'istante "1" e per tutti gli istanti successivi.
$\Rightarrow$ La sorgente è quindi **stazionaria** e W viene detta **distribuzione stazionaria di S**.

N.B.:
Ogni sorgente con memoria finita ha almeno una distribuzione stazionaria. $\rightarrow$ Ma in generale non sappiamo se $\exists$ e se è unica

Teorema 2.3

Le distribuzioni stazionarie di una sorgente con matrice di Markov Π sono tutti e soli i vettori di prob.tà nello spazio generato dagli autovettori sinistri relativi agli autovalori di Π uguali ad "1".

OSS: In un vett. di prob.tà tutti gli elem. sono ≤ 1 e la Σ è 1

OSS: Se la sorgente ha mem. finita esiste sempre la corrispond. catena di Markov.

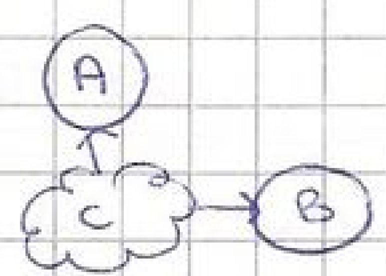
OSS: Le distrib. staz. quindi sono tutti e soli i vett. di prob.tà che sono anche autovettori di Π (relativi a $\lambda = 1$) ok vero??

N.B.:
NON basta che sia "autovett.", deve anche essere un vett. di prob.tà, cioè con tutti gli elementi ≥ 0.

Sorgenti decomponibili:

Def: Un insieme B di stati t.c. ogni elem. di B è raggiungibile da ogni altro elem. di B e che NESSUN elemento fuori di B è raggiungibile da un elemento di B è detto insieme essenziale.

x es.:



A e B sono essenziali, ma se ne può uscire. C è un insieme di stati transitori.

Def: Una sorgente è detta indecomponibile se ha un solo insieme essenziale.
È detta decomponibile se ne ha più di uno.

OSS: Se la matrice di Markov Π è IRRIDUCIBILE vuol dire che da ogni stato si può andare in un qualunque altro stato in un n° finito di passi e con probab. NON nulla.

$\Rightarrow$ Tutti gli stati della sorgente formano un solo insieme essenziale, ossia S è indecomponibile.

Teorema:

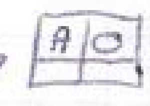
Se la matrice di Markov è IRRIDUCIBILE: forse non ci vuole perché sembra ∄ matrici IRR. e NON regolari!!

1) ∃ un solo autovalore = 1 (e tutti gli altri in modulo sono strettamente < 1) e il corrisp. autovett. x è tutto positivo.

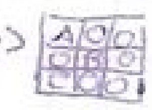
2) $\Rightarrow$ ∃ un solo autovettore associato ad "1": $X^T \Pi = X^T$ con $X^T > 0$
cioè S ammette un'unica distrib. stazionaria.

Riassunto

Matrice	Autovalori uguali a 1	Insiemi essenziali	Sorgente	Distribuzioni stazionarie
	Se:	Se:	Se:	allora:
Irriducibile	1	1	Indecomponibile	1
Riducibile	1	1	Indecomponibile	1
Riducibile	*k*	*k*	Decomponibile	∞



→ 1a sola matr. IRR



→ Rid. ma con + sottomatrici IRR

→ A tutti gli effetti ho 2 o + sorgenti * e qualunque comb. lin. delle loro distr. staz. è ancora una distr. staz., quindi sono ∞.

Schemi di decisione e probabilità di errore

Def: Dato un canale con r simboli in ingresso uno **schema di decisione** è una partizione dell'insieme dei simboli di uscita B in r sottoinsiemi disgiunti $B_1, \dots, B_r$, di cui ciascuno può essere vuoto. Più formalmente vale:

$$\boxed{B_u \cap B_v = \emptyset \text{ per } u \neq v \text{ e } B = \bigcup_{i=1}^{r} B_i}$$

Consideriamo quindi una funzione **dec** da $\{1, \dots, s\}$ in $\{1, \dots, r\}$ tale che:

$dec(j)$ è l'unico i per cui $b_j \in B_i$

(con $j = 1, \dots, s$ e $i = 1, \dots, r$)

OSS: si assume $r \leq s$, perché non ha senso il contrario (alcuni simboli in ingresso NON potrebbero essere mai decodificati).

Probabilità di errore

- $\boxed{P_{ok}(b_j) = p(a_{dec(j)} \mid b_j)}$ → prob.tà di azzeccarci dopo aver ricevuto b_j
- $\boxed{p_{err} = \sum_{j=1}^{s} p(b_j) \cdot (1 - p(a_{dec(j)} \mid b_j))}$ → prob.tà totale di errore per un certo schema di decisione, si basa su E/U ⊛
- $\boxed{p_{err}(a_i) = \sum_{b_j \notin B_i} p(b_j \mid a_i)}$ → prob.tà di errore dato un simbolo a_i
- $\boxed{p_{err} = \sum_{i=1}^{r} p(a_i) \sum_{b_j \notin B_i} p(b_j \mid a_i)}$ → prob.tà totale di errore in funzione di $p_{err}(a_i)$, si basa su U/E

La formula da utilizzare si sceglie in base alla matrice che abbiamo: U/E o E/U

Schema dell'osservatore ideale:

Si sceglie la funz. di decodifica dec(j) in modo che:

$$\boxed{p(a_{dec(j)} \mid b_j) = \max_i \; p(a_i \mid b_j)}$$

In questo modo p_{err} è minimo, infatti ogni termine della ⊛ è minimizzato.

Schema della massima verosimiglianza:

Se la distrib. di entrata ha tutti i simboli equiprobabili si ha:

$$p(a_i \mid b_j) = \frac{p(b_j \mid a_i)}{r \cdot p(b_j)}$$

e trovare il max su i delle prob.tà all'indietro equivale a massimizzare le prob.tà in avanti, ovvero trovare il valore dec(j) tale che:

$$\boxed{p(a_{dec(j)} \mid b_j) = \frac{1}{r\, p(b_j)} \max_i \; p(b_j \mid a_i)}$$

Tale schema si può usare quando non si conosce la distribuzione di entrata.

OSS: Se una sorg. S è compressa bene, quello che arriva al canale sono bit equiprobabili e indip.. In tal caso l'os. ideale e la max verosim. coincidono e quindi la distrib. di entrata non ci interessa.

Codifica del canale in presenza di Rumore.

Velocità di trasmissione:

Consideriamo un canale con:

- alfabeto di ingresso con r simboli
- alfabeto di uscita con s simboli
- capacità C

Supp. di dover usare il canale una volta al secondo, il passaggio di info attraverso il canale è al più:

$$\boxed{\log r \text{ bit/sec}}$$

Se consideriamo come alfabeto di ingresso tutte le poss. stringhe lunghe "n" queste saranno "r^n".

$\Rightarrow$ la velocità di trasmissione è: $\boxed{\dfrac{\log(r^n)}{n} = \log r}$

Se però di queste r^n stringhe ne usiamo solo q ($\leq r^n$):

$$\boxed{R = \frac{\log q}{n} \text{ bit/s}}$$

Teorema Fondamentale (4.1)

Dato un canale con r simboli in ingresso,
s simboli in uscita,
capacità C.
Per ogni $R < C$ è possibile trovare una sequenza di codici di lunghezza crescente n con velocità di trasmissione R e con prob.tà di errore che tende a 0.

Ossia, ci possiamo riuscire purchè la velocità di trasm. (R) sia inferiore alla capacità del canale (C).

Lemma 4.1:

Per ogni codice di q parole ed ogni schema associato ad esso, se le parole del codice sono fatte entrare nel canale con distrib. di prob.tà E e si considera la distrib. delle parole di uscita U e la prob.tà di errore per vale:

$$\boxed{H(E|U) \leq H(p_{er}, 1-p_{er}) + p_{er} \cdot \log(q-1)}$$

Lemma 4.2:

Si consideri una sequenza $x_1, x_2, \dots, x_n$ di simboli in ingresso ad un canale e la seq. dei corrisp. $y_1, \dots, y_n$ simboli di uscita, vale:

$$\boxed{I(x_1,\dots,x_n | y_1,\dots,y_n) \leq \sum_{i=1}^{n} I(x_i | y_i) \leq nC}$$

Lemma 4.3:

La prob.tà di errore per di un qualunque codice con q parole di lunghezza n, usato per un canale di capacità C deve soddisfare la relazione:

$$\boxed{\log q \leq \frac{nC+1}{1-p_{er}}}$$

Teorema 4.2: (Inverso debole)

1) Per ogni n finito $p_{er} = 0$ implica $R \leq C$

2) Se $R > C$ non $\exists$ una seq. di codici con p_{er} che tende a 0 per n che tende all'infinito.

ESERCIZI

Entropia:

(1) Sappiamo che: $H(p_1, \dots, p_K) = -\sum_{i=1}^{K} p_i \log p_i$ con $\sum_{i=1}^{K} p_i = 1$

Calcolare:

$$\underline{H(2^{-1}, 2^{-2}, \dots, 2^{-K}, 2^{-K})}$$

$$= -\sum_{i=1}^{K} 2^{-i} \log 2^{-i} + (-2^{-K} \log 2^{-K}) =$$

$$= -\sum_{i=1}^{K} 2^{-i} \cdot (-i) - 2^{-K} \cdot (-K) =$$

$$= \sum_{i=1}^{K} i \cdot 2^{-i} + K \cdot 2^{-K}$$

Calcolare con: $K = 4, 6, 8$ e 100

con $K = 4$:

$$1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + 4 \cdot \frac{1}{16} = \frac{8}{16} + \frac{8}{16} + \frac{6}{16} + \frac{4}{16} + \frac{4}{16} = \frac{30}{16}$$

con $K = 6$:

...

⇒ In pratica si ricava che la formula generale è:

$$\boxed{\frac{(2 \cdot 2^i) - 2}{2^i}}$$

Classificazione delle sorgenti

① (2.6)

Calcolare la sorgente adiacente della sorg. def. dalla matrice:

	0	1
0	1	0
1	0,6	0,4

Gli autovalori di questa matrice sono 1 e 0,4,

$\Rightarrow$ $\exists$ un solo autovalore $= 1$ $\Rightarrow$ la sorgente è regolare.

$\Rightarrow$ $\exists$ una sola distrib. stazionaria che è data dal vettore x^T tale che $x^T = x^T \cdot \begin{pmatrix} 1 & 0 \\ 0,6 & 0,4 \end{pmatrix}$

Risolvo il sistema:

$$\begin{cases} x_1 + 0,6\, x_2 = x_1 \\ 0,4\, x_2 = x_2 \end{cases}$$

Il determinante di questo sistema è $\emptyset$, quindi ci sono ∞ soluzioni.

$\Rightarrow$ ∞ autovettori e UNA sola distrib. stazionaria

La distribuzione stazionaria si trova normalizzando in norma 1, ossia la Σ degli elementi del vettore deve essere 1.

$$\begin{cases} x_1 = \text{qualsiasi} \\ x_2 = \emptyset \end{cases} \Rightarrow x^T = (1, 0)$$

2) Calcolare la sorgente adiacente della sorgente della matrice:

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 2/3 & 1/3 \\ 1 & 1/3 & 2/3 \end{array}$$

Come già visto dobbiamo risolvere:

$$x^T \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix} = x^T$$

$$\Downarrow$$

$$(x_1, x_2) \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}$$

Ci aspettiamo un sistema con determinante $= \emptyset$ ($\emptyset$ non $+\infty$) e quindi ∞ soluzioni ($\det = \emptyset$ significa che ci sono almeno 2 righe lin. dip.)

$$\Rightarrow \begin{cases} 2/3\, x_1 + 1/3\, x_2 = x_1 \\ 1/3\, x_1 + 2/3\, x_2 = x_2 \end{cases} \Rightarrow \begin{cases} 2x_1 + x_2 = 3x_1 \\ x_1 + 2x_2 = 3x_2 \end{cases} \Rightarrow \begin{cases} -x_1 = -x_2 \\ -x_2 = -x_1 \end{cases}$$

Quando ci sono più equazioni uguali è suff. tenerne una:

$\Rightarrow \quad x_1 = x_2$, normalizziamo in norma 1 $\Rightarrow x^T = (1/2, 1/2)$

OSS: si vedeva ad occhio perché la matrice è simmetrica.

3) Calcolare la sorgente adiacente della sorg. dip. dalla matrice:

	0	1
0	1/2	1/2
1	1/2	1/2

questa sorgente ha $m = 0$, infatti le righe di prob.tà sono le stesse.

$\Rightarrow$ la sorgente coincide con la propria adiacente e la distribuzione stazionaria è $X^T = (1/2, 1/2)$
(la matrice è anche simmetrica) OK?? è per questo??

Calcolare l'estensione seconda:

S^2 emetterà le seguenti coppie di simboli:
00, 01, 10, 11 con prob.tà 1/4, 1/4, 1/4, 1/4, infatti S è senza memoria.

	00	01	10	11
00	1/2	1/2	0	0
01	0	0	1/2	1/2
10	1/2	1/2	0	0
11	0	0	1/2	1/2

$\rightarrow$ fatta io, OK??

Ovviamente S e S^2 non coincidono, la prima ha 2 simboli e la seconda ne ha 4.

4) Calcolare l'estensione seconda della sorgente definita dalla matrice:

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 1/2 & 1/2 \end{array}$$

$S^2 = \{00, 01, 10, 11\}$, S ha $m = 1$

$\Rightarrow$

	00	01	10	11
00	$1 \cdot 1$	$\emptyset$	$\emptyset$	$\emptyset$
01	$\frac{1}{2} \cdot 1$	$\frac{1}{2} \cdot \emptyset$	$\frac{1}{2} \cdot \frac{1}{2}$	$\frac{1}{2} \frac{1}{2}$
10	$1 \cdot 1$	$1 \cdot 0$	$0 \cdot \frac{1}{2}$	$0 \cdot \frac{1}{2}$
11	$\frac{1}{2} \cdot 1$	$\frac{1}{2} \cdot 0$	$\frac{1}{2} \cdot \frac{1}{2}$	$\frac{1}{2} \cdot \frac{1}{2}$

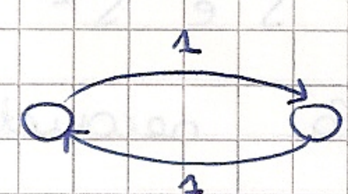
5) Dare un esempio di S:
- NON regolare
- indecomponibile
- con 7 simboli

Se non viene data si assume $m = 1$.

Conviene rappr. con il grafo, dove è più semplice evidenziare le proprietà.

con $m = 1$ per 7 simboli abbiamo 7 stati.
(con $m = 2$ avremmo avuto 7^2 stati!)

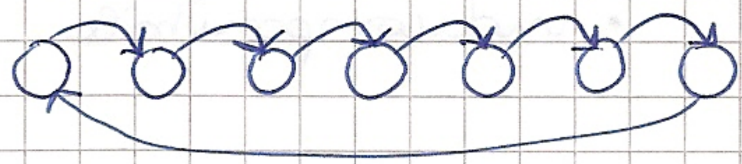
Partiamo da un caso base:



2 stati, da uno passo all'altro $\Rightarrow$ ho un solo insieme essenziale $\Rightarrow$ INDECOMPONIBILE

Affinché sia NON REGOLARE devono esserci più autovalori in modulo $= 1$ $(1, -1)$ (se le prob-tà erano $1/3$ e $2/3$ era cmq NON regolare perché non converge, NON si finisce in uno stato per restarci)

Estendiamo a 7 simboli:



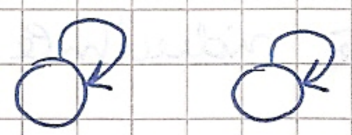
Il fatto che sia indecomp. si vede anche dal fatto che il grafo è fortemente connesso.

6) Dare un esempio di S:
- decomponibile
- NON regolare

(se è decomp. è sempre NON regolare)

Deve essere il più semplice possibile:

Per essere decomp. deve avere almeno 2 insiemi essenziali:



La matrice che corrisponde a questo grafo è l'identica 2×2 che tra l'altro è riducibile perché già ridotta.
L'identica, in generale, ha tanti insiemi essenziali quante sono le sue righe, quindi è utile nel caso si dovessero dare un certo n° di simboli.

7) Dare un esempio di S: • regolare
• indecomponibile

Se la vogliamo REGOLARE va fatta in modo che sia evidente che c'è un solo autovalore = 1.

Se NON è specificato si fa con 2 simboli.

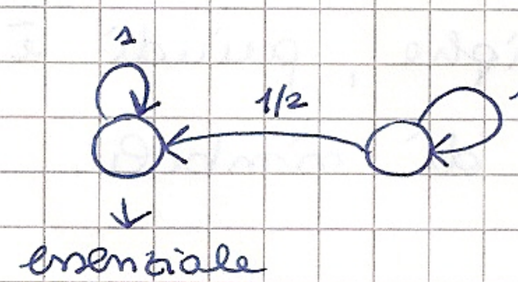
La matrice non può essere diagonale, perché altrimenti verrebbe l'identica, con due autovalori in modulo = 1 e due insiemi essenziali, sarebbe NON regolare e decomponibile.

Quindi per scegliere gli autovalori a nostro piacere la facciamo triangolare:

1	0
1/2	1/2

→ questa matrice è riducibile, ma non ci interessa

Che ha un solo insieme essenziale si può osservare anche dal grafo:



Se l'avessimo voluta con 3 simboli:

1	0	0
1/2	1/2	0
1/3	1/3	1/3

→ questo tipo di S si ricava facilmente con qualsiasi n° di simboli.

OSS: Anche dal grafo si vede che è regolare, infatti c'è convergenza, prima o poi si finisce nel 1° stato e ci si resta. La distr. staz. è (1,0) infatti all'infinito con prob.tà 1 vado nel primo stato.

Algoritmo di Huffman

1) Supp. di avere una S che emette:

a	con	prob.tà	0,45
b	"	"	0,2
c	"	"	0,2
d	"	"	0,1
e	"	"	0,05

la Σ deve essere ovviamente 1

Il codice migliore deve dare parole più lunghe alle uscite meno probabili e più corte a quelle meno probabili.

Si parte considerando i 2 eventi meno probabili e poi si assegnano 2 bit diversi:

$\Rightarrow$

a	0,45	
b	0,2	
c	0,2	
d	0,1	0
e	0,05	1

la nuova S è:

a	0,45
b	0,2
c	0,2
d, e	0,15 = 0,1 + 0,05

Iteriamo:

$\Rightarrow$

a	0,4	
b	0,2	
c	0,2	0
d, e	0,15	1

→ NON è univoco, potevamo prendere anche "b" che ha la stessa prob.tà di "c"

$\longrightarrow$

a	0,4
b	0,2
c, d, e	0,35 = 0,2 + 0,15

Riordiniamo:

a	0,4	
c, d, e	0,35	0
b	0,2	1

$\longrightarrow$

a	0,4	0
b, c, d, e	0,55 = 0,35 + 0,2	1

In pratica dobbiamo ricavato un albero che ci da il codice istantaneo per S, basta percorrerlo a ritroso:

$a = 0$

$b = 11$

$c = 100$

$d = 1010$

$e = 1011$

2) Data S: a con prob.tà 0,2
b " " 0,4
c " " 0,1
d " " 0,3

Calcolare il codice istantaneo:

b	0,4	
d	0,3	
a	0,2	0
c	0,1	1

→

b	0,4	
d	0,3	0
a,c	0,3	1

→

b	0,4	1
a,c,d	0,6	0

$\Rightarrow$ $a = 010$

$b = 1$

$c = 011$

$d = 00$

Calcoliamo $H(S)$ e $\bar{n}$:

$H(S) = -(0{,}2 \log 0{,}2 + 0{,}3 \log 0{,}3 + 0{,}2 \log 0{,}2 + 0{,}1 \log 0{,}1) = 1{,}846$

$\bar{n} = 0{,}4 \cdot 1 + 0{,}3 \cdot 2 + 0{,}2 \cdot 3 + 0{,}1 \cdot 3 = 1{,}9$

$H(S) \neq \bar{n}$ non è ottimo, $\bar{n}$ si avvicina ad $H(S)$, ma non potrà mai essere più piccola..

3) Data S : a con prob.tà 1/2

b " " 1/4

c " " 1/8

d " " 1/8

Trovare : $H(S)$, codice istantaneo, $\overline{n}$ (lunghezza media del codice $= \sum_{i=1}^{k} p_i |x_i|$), $p(0)$, $p(1)$

Codice istantaneo:

a	1/2		→	a	1/2		→	a	1/2	1
b	1/4			b	1/4	1		b,c,d	1/2	0
c	1/8	1		c,d	1/4	0				
d	1/8	0								

⇒ a = 1, b = 01, c = 001, d = 000

$$H(S) = -\left(\frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4} + \frac{1}{8}\log\frac{1}{8} + \frac{1}{8}\log\frac{1}{8}\right) = \frac{7}{4}$$

$$\overline{n} = \frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{8}\cdot 3 + \frac{1}{8}\cdot 3 = \frac{7}{4}$$

$H(S) = \overline{n}$ ⇒ abbiamo raggiunto l'ottimo → Vedere perché!!!

⇒ $p(0)$ e $p(1)$ dovranno essere $\frac{1}{2}$ e $\frac{1}{2}$

$$p(1) = \frac{1/2\cdot 1 + 1/4\cdot 1 + 1/8\cdot 1 + 1/8\cdot 0}{7/4} = \frac{1}{2}$$

$$p(0) = \frac{1/2\cdot 0 + 1/2\cdot 1 + 1/8\cdot 2 + 1/8\cdot 3}{7/4} = \frac{1}{2}$$

4) Data S:

a	0,25
b	0,25
c	0,25
d	0,25

Trovare: $H(S)$, codice istantaneo, $\bar{n}$, $p(0)$, $p(1)$

Codice istantaneo:

a	0,25		→	c,d	0,50		→	c,d	0,50	0
b	0,25			a	0,25	0		a,b	0,50	1
c	0,25	0		b	0,25	1				
d	0,25	1								

⇒ a = 10, b = 11, c = 00, d = 01

$H(S) = 2$, l'entropia di un exp. equiprob. è $\log k$ n° uscite

$\bar{n} = 2$ ovviamente

⇒ abbiamo raggiunto l'ottimo ⇒ $p(0) = 1/2$, $p(1) = 1/2$

Cambiamo di poco le probabilità e vediamo che succede:

a	0,27		→	a	0,27	1	→	a,b	0,53	1	⇒	a	= 11
b	0,26			b	0,26	0		c,d	0,47	0		b	= 10
c	0,24	1		c,d	0,47							c	= 01
d	0,23	0										d	= 00

$H(S) = 1,997$ → quando l'evento è sbilanciato l'info. associata diminuisce, quindi anche l'entropia

$\bar{n} = 2$

$p(0) = 0,48$

$p(1) = 0,52$

2007-0 / 2009-3 1)

ENTROPIA

1) Determinare l'entropia dell'esperimento con probabilità $(1/128, 1/128, 1/128, 1/128, 31/32)$.

L'entropia di un esperimento X si calcola con la formula

$$H(X) = -C\sum_{i=1}^{k} p_i \log p_i \quad \text{dove} \quad C = \frac{1}{\log_2 2} = 1$$

$$\Rightarrow H(X) = -\left(\frac{1}{128}\log\frac{1}{128} + \frac{1}{128}\log\frac{1}{128} + \frac{1}{128}\log\frac{1}{128} + \frac{1}{128}\log\frac{1}{128} + \frac{31}{32}\log\frac{31}{32}\right) =$$

$$= -\left(\frac{4}{128}\log\frac{1}{128} + \frac{31}{32}\log\frac{31}{32}\right) =$$

$$= -\left(-\frac{28^{7}}{\cancel{128}\,32} + \frac{31}{32}(\log 31 - 5)\right) =$$

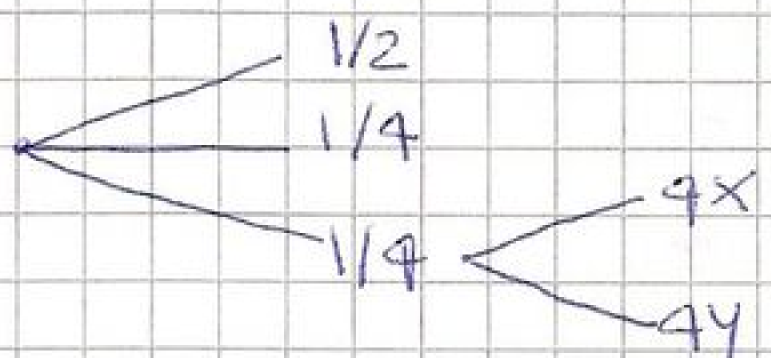
$$= -\left(-\frac{7}{32} - \frac{155}{32} + \frac{31}{32}\log 31\right) = \frac{162}{32} - \frac{31}{32}\log 31 =$$

$$= \frac{81}{16} - \frac{31}{32}\log 31 = 0{,}26\ldots$$

2007-2 / 2008-1 1)

1) Determinare due valori di x e y per cui $H(2^{-1}, 2^{-2}, x, y)$ sia minimo, giustificare il risultato.

Possiamo considerare l'esperimento come composto:



dove $x + y = \frac{1}{4}$

e $4x + 4y = 1$

Per il terzo assioma dell'entropia, l'entropia risulta essere:

$$H(X) = H(1/2, 1/4, 1/4) + \frac{1}{4}H(4x + 4y)$$

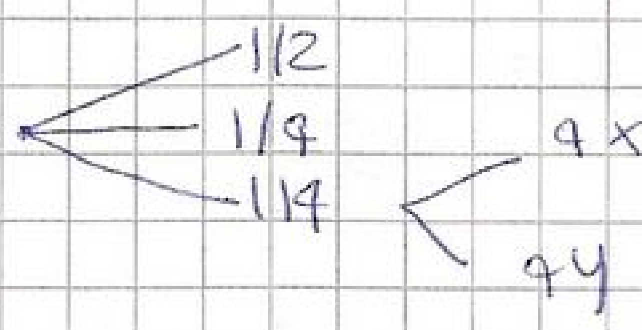
Per minimizzare è sufficiente scegliere o x o y uguale a $\emptyset$.

$$\Rightarrow H(X) = H(1/2, 1/4, 1/4) = -\left(-\frac{1}{2} - 1\right) = \frac{3}{2}$$

2007-3 | 2008-3 1)

1) Determinare due valori di x e y per cui $H(2^{-1}, 2^{-2}, x, y)$ sia massimo.

Possiamo vedere l'esperimento come composto:



con: $x + y = 1/4$

$4x + 4y = 1$

Per il terzo assioma dell'entropia, l'entropia risulta essere:

$$H(X) = H(1/2, 1/4, 1/4) + \frac{1}{4} H(4x, 4y)$$

Per massimizzare è suff. scegliere $x = y$ in questo caso $x = y = \frac{1}{8}$

$$\Rightarrow H(X) = H(1/2, 1/4, 1/4) + \frac{1}{4} H(1/2, 1/2) = \frac{3}{2} + \frac{1}{4} = \frac{7}{4}$$

2007-4 / 2008-2 1) / 2009-2 1)

1) Determinare due valori di x e y per cui $H(2^{-1}, 2^{-2}, x, y)$ sia un numero irrazionale.

Basta scegliere x e y in modo che non sia una potenza di 2 e che $x + y = \frac{1}{4}$.

$$x = \frac{1}{9} \qquad y = \frac{1}{4} - \frac{1}{9} = \frac{9-4}{36} = \frac{5}{36}$$

$$H(X) = -\left(\frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4} + \frac{1}{9}\log\frac{1}{9} + \frac{5}{36}\log\frac{5}{36}\right) =$$

$$= -\left(-\frac{1}{2} - \frac{1}{2} + \frac{1}{9}\log\frac{1}{9} + \frac{5}{36}\log\frac{5}{36}\right) = 1{,}74776\ldots$$

2007 - 5 / 2008 - 0 1) / 2008 - 4 1)

1) Determinare due valori di x e y per cui $H(2^{-1}, 2^{-2}, x, y)$ sia un numero razionale.

È sufficiente prendere x e y in modo che siano potenze di 2, per esempio $x = y = \frac{1}{8}$ con $x + y = \frac{1}{4}$.

2007 - 6

1) Determinare l'entropia dell'esperimento con prob.tà $(1/128, 1/128, 1/128, 1/128, 31/32)$

$$H(X) = -\sum_{i=1}^{K} p_i \log p_i$$

$$H(X) = -\left(\frac{4}{128}\log\frac{1}{128} + \frac{31}{32}\log\frac{31}{32}\right) = -\left(-\frac{7}{32} - \frac{5}{32} + \frac{31}{32}\log 31\right) =$$

$$= \frac{162}{32} - \frac{31}{32}\log 31 = 0{,}2631\ldots$$

2008 - 5

1) Determinare l'entropia dell'esperimento con prob.tà $(1/2, 1/4, 1/8, 2^{-k+1}, 2^{-k}, 2^{-k})$

Deve essere: $2^{-k+1} + 2^{-k} + 2^{-k} = \frac{1}{8}$

$$\Rightarrow 2^{-k+1} + 2 \cdot 2^{-k} = 2^{-k+1} + 2^{-k+1} = 2 \cdot 2^{-k+1} = 2^{-k+2}$$

$$\Rightarrow 2^{-k+2} = 2^{-3} \Rightarrow -k+2 = -3 \Rightarrow \cancel{-}k = \cancel{-}5$$

$$H(1/2, 1/4, 1/8, 1/16, 1/32, 1/32) =$$

$$= -\left(1/2 \log 1/2 + 1/4 \log 1/4 + \tfrac{1}{8}\log\tfrac{1}{8} + 1/16 \log 1/16 + 2/32 \log 1/32\right) =$$

$$= -\left(-\frac{1}{2} - \frac{1}{2} - \frac{3}{8} - \frac{1}{4} - \frac{5}{16}\right) = -\left(\frac{-8-8-4-5-6}{16}\right) = \frac{31}{16}$$

2007-2

ESEMPI DI SORGENTI

2) Dare un esempio di sorgente regolare con memoria 1 e almeno 10 simboli.

	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10
q1	1									
q2	1/2	1/2								
q3	1/3	1/3	1/3					$\emptyset$		
q4	1/4	1/4	1/4	1/4						
q5	1/5	1/5	1/5	1/5	1/5					
q6										
q7										
q8										
q9										
q10	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10

È regolare perché ha un solo autovalore uguale a 1, ha memoria 1 poiché viene considerato un solo simbolo per ogni stato.

2007-6 / 2008-2 2) / 2008-2 2)

2) Dare un esempio di sorgente indecomponibile, con matrice associata irriducibile, non regolare e uno di sorgente indecomponibile, regolare, con matrice associata riducibile

1) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

indecomp: un solo ins. erg.

irrid.: da ogni stato raggiungo ogni altro stato

NON regolare: ha due autovalori in modulo uguali a 1

2) $\begin{pmatrix} 1 & 0 \\ 1/2 & 1/2 \end{pmatrix}$

indecomp.: un solo ins. erg.

rid.: già ridotta

regolare: c'è un solo autovalore uguale a 1.

2008-0

2) Dare un esempio di sorgente regolare con memoria 1 e almeno 3 simboli.

	a_1	a_2	a_3
a_1	1	0	0
a_2	1/2	1/2	0
a_3	1/3	1/3	1/3

Regolare – ha un solo autovalore uguale a 1

3 simboli – $\{a_1, a_2, a_3\}$

$m = 1$ – ogni stato è composto da un solo simbolo

2007-0 ESTENSIONI - ADIACENTI

2) Calcolare l'estensione seconda della sorgente definita dalla matrice:

	0	1
0	0	1
1	1	0

Dire che memoria ha la sorgente risultante.

$S^2 = \{00, 01, 10, 11\}$ l'estensione seconda ha 4 simboli

	00	01	10	11
00	0	0	1	0
01	0	1	0	0
10	0	0	1	0
11	0	1	0	0

S^2 ha memoria 1 poiché la memoria dell'estensione si calcola come $\left\lceil \frac{m}{k} \right\rceil$ in questo caso $\left\lceil \frac{1}{2} \right\rceil = 1$

2007-1 / 2008-6 2)

2) Calcolare la sorgente adiacente della sorgente definita dalla matrice:

	0	1
0	0	1
1	1	0

La sorgente $\bar{S}$ è senza memoria rispetto ad S, troviamo una distribuzione stazionaria (S è NON regolare) risolvendo il sistema $\bar{x}A = \bar{x}$

$$\begin{cases} x_2 = x_1 \\ x_1 = x_2 \end{cases}$$

$\Rightarrow$ Normalizzo in norma 1

$\Rightarrow x^T = (1/2, 1/2)$

2007-3

2) Calcolare la sorgente estensione seconda della sorgente definita dalla matrice:

	0	1
0	a	1-a
1	a	1-a

Determinare la memoria della sorgente risultante.

$S^2 = \{00, 01, 10, 11\}$ (tutte le possibili combinazioni dei simboli di S)

	00	01	10	11
00	a^2	$a(1-a)$	$(1-a)a$	$(1-a)(1-a)$
01	$a \cdot a$	$a(1-a)$	$(1-a)a$	$(1-a)(1-a)$
10	a^2	$a(1-a)$	$(1-a)a$	$(1-a)(1-a)$
11	a^2	$a(1-a)$	$(1-a)a$	$(1-a)(1-a)$

S^2 ha memoria $\left\lceil \frac{m}{k} \right\rceil = \left\lceil \frac{1}{2} \right\rceil = 1$

l'estensione si comporta esattamente come S, ma le uscite sono raggruppate.

2007 - 4 / 2008 - 4 2)

2) Calcolare la sorgente adiacente della sorgente def. dalla matrice:

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & a & 1-a \\ 1 & 1-b & b \end{array}$$

La sorgente $\bar{S}$ è una semplificazione di S poiché è senza memoria, quindi per trovare la distrib. stazionaria dobbiamo risolvere il sistema: $x^T A = x^T$

$$(x_1, x_2) \cdot \begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix} = (x_1, x_2)$$

$$\Rightarrow \begin{cases} a x_1 + (1-b) x_2 = x_1 \\ (1-a) x_1 + b x_2 = x_2 \end{cases} \qquad \begin{cases} (a-1) x_1 = (-1+b) x_2 \\ \cancel{(b-1) x_2 = (a-1) x_1} \end{cases}$$

$$x_1 = \frac{(b-1)}{(a-1)} x_2 \qquad x^T = \left(\frac{b-1}{a-1} x_2 , x_2 \right)$$

normalizzo in norma 1 : $\dfrac{b-1}{a-1} x_2 + x_2 = 1$

$$\Rightarrow \left(\frac{b-1}{a-1} + 1 \right) x_2 = 1 \quad \Rightarrow \quad x_2 = \frac{1}{\frac{b-1}{a-1} + 1}$$

$$\Rightarrow x^T = \left(\frac{\frac{b-1}{a-1}}{\frac{b-1}{a-1} + 1} , \frac{1}{\frac{b-1}{a-1} + 1} \right)$$

2008-3

2) Calcolare la sorgente estensione seconda della sorgente definita dalla matrice:

	0	1
0	a	$1-a$
1	$1-b$	b

$S^2 = \{00, 01, 10, 11\}$ S^2 ha 4 simboli

	00	01	10	11
00	$a \cdot a$	$a(1-a)$	$(1-a)(1-b)$	$(1-a)b$
01	$(1-b)a$	$(1-b)(1-a)$	$b(1-b)$	$b \cdot b$
10	a^2	$a(1-a)$	$(1-a)(1-b)$	$(1-a)b$
1	$(1-b)a$	$(1-b)(1-a)$	$b(1-b)$	b^2

2008-3

2) Calcolare la sorgente estensione seconda della sorgente definita dalla matrice:

	0	1
0	a	$1-a$
1	$1-a$	a

$S^2 = \{00, 01, 10, 11\}$ S^2 ha 4 simboli

	00	01	10	11
00	a^2	$a(1-a)$	$(1-a)(1-a)$	$(1-a)a$
01	$(1-a)a$	$(1-a)(1-a)$	$a(1-a)$	a^2
10	a^2	$a(1-a)$	$(1-a)(1-a)$	$(1-a)a$
11	$(1-a)a$	$(1-a)(1-a)$	$a(1-a)$	a^2

2008 - 4

2) Calcolare la sorgente adiacente della sorgente definita dalla matrice :

	0	1
0	a	$1-a$
1	$1-a$	a

$\overline{S}$ è una semplificazione di S, ossia è senza memoria, per trovare la distribuzione stazionaria bisogna risolvere il sistema : $x^T A = x^T$

$$(x_1, x_2) \begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix} = (x_1, x_2)$$

$$\Rightarrow \begin{cases} a x_1 + (1-a) x_2 = x_1 \\ (1-a) x_1 + a x_2 = x_2 \end{cases} \Rightarrow \begin{cases} (a-1) x_1 = (a-1) x_2 \\ \cancel{(a-1) x_2 = (a-1) x_1} \end{cases}$$

$x_1 = x_2$ normalizzo in norma 1 $x^T = (1/2, 1/2)$

2007-3

CODICI UNIVOC. DECIFRABILI / ISTANTANEI

3) Dare un esempio di codice non univocamente decifrabile nell'alfabeto {a, b, c}.

a
aa
ab
ac

non è univoc. decifrabile perchè se riceviamo "aa" non si sa se è una sola parola di codice (aa) o due in sequenze (a+a).

2007-5 / 2008-0 3)

3) Dare un esempio di codice univocamente decifrabile non istantaneo nell'alfabeto {a, b, c}

a
ab
ac

è univoc. decifrabile perchè ad ogni parola del codice può essere associata una sola seq. di simboli della sorgente, è non istantaneo poichè "a" è prefisso di "ab" e "ac".

2008-3

3) Dare un esempio di codice non univocamente decifrabile nell'alfabeto {0, 1}.

C = { 0, 00, 01 }

Non è univocamente decifrabile perchè se riceviamo "00" non sappiamo se si tratta del simbolo "00" o dei simboli "0" + "0".

2009-3

3) Dare un esempio di codice istantaneo nell'alfabeto $\{0,1,2,3\}$

$C = \{00, 11, 22, 33\}$

è istantaneo poiché le parole hanno tutte la stessa lunghezza, ciò equivale a scegliere fra le foglie di lunghezza 2, nessuna parola è prefisso di un'altra.

2007 - 0

HUFFMAN

3) Dato il codice binario di Huffman per una sorgente senza memoria con simboli $\{a, b, c, d\}$ e con probabilità $\{0{,}9, 0{,}09, 0{,}009, 0{,}001\}$. Calcolare la prob.tà che dopo la codifica venga almeno uno 0.

a	0,9		a	0,9		a	0,9	0
b	0,09		b	0,09	0	b,c,d	0,1	1
c	0,009	0	c,d	0,01	1			
d	0,001	1						

$a = 0$, $b = 10$, $c = 110$, $d = 111$

Calcolo la lunghezza media:

$$\bar{n} = \sum_{i=1}^{K} p_i |x_i| = 0{,}9 \cdot 1 + 0{,}09 \cdot 2 + 0{,}009 \cdot 3 + 0{,}001 \cdot 3 =$$
$$= 0{,}9 + 0{,}18 + 0{,}027 + 0{,}003 = 1.11$$

$$p(0) = \frac{0{,}9 \cdot 1 + 0{,}09 \cdot 1 + 0{,}009 \cdot 1}{1{,}11} = \frac{0{,}999}{1{,}11} = 0{,}9$$

$\Rightarrow p(1) = 0{,}1$

2007-1 **CANALI**

3) Dare un esempio di canale con capacità 7.

In un canale lossless $C = \log r$, dove r è il numero di simboli in entrata e vi è un solo elemento $\neq \emptyset$ per ogni colonna

$$\begin{array}{c|ccccccc}
 & b_1 & b_2 & b_3 & \cdots & b_{127} & b_{128} & b_{129} \\
\hline
a_1 & 1 & & & & & & \\
a_2 & & 1 & & & \emptyset & & \\
a_3 & & & 1 & & & & \\
\vdots & & & & \ddots & & & \\
 & & \emptyset & & & & & \\
a_{127} & & & & & 1 & & \\
a_{128} & & & & & & 1/2 & 1/2
\end{array}$$

2007-2 | 2008-1 3)

3) Dare un esempio di canale con capacità $\emptyset$.

Il canale BSC:

	0	1
0	1/2	1/2
1	1/2	1/2

ha capacità $\emptyset$, in pratica non passa alcuna informazione.
La capacità nei BSC si calcola come:

$$C = 1 - H(p,q) \qquad \text{con } p = q = 1/2 \Rightarrow H(p,q) = 1$$

$$\Rightarrow C = \emptyset$$

2007-4 / 2008-4 3)

3) Dare un esempio di canale lossless con capacità 3 e uno di canale deterministico con capacità 2:

lossless:

$C = \log r$, r n° di simboli in entrata, un solo elem. $\neq \emptyset$ per colonna $\Rightarrow$ $r = 8$

	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
a_1	1								
a_2		1							
a_3			1				$\emptyset$		
a_4				1					
a_5					1				
a_6		$\emptyset$				1			
a_7							1		
a_8								1/2	1/2

Deterministico:

$C = \log s'$, s' n° di simboli in uscita realmente usati, un solo elem. $\neq 0$ per riga $\Rightarrow$ $s' = 4$

	b_1	b_2	b_3	b_4	b_5
a_1	1	0	0	0	0
a_2	0	0	1	0	0
a_3	0	0	0	1	0
a_4	0	0	0	0	1

2008 - 2

3) Dato il canale

$$\begin{array}{c|cc} & b_1 & b_2 \\ \hline a_1 & 0{,}7 & 0{,}3 \\ a_2 & 0{,}3 & 0{,}7 \end{array}$$

enunciare tutti i possibili schemi di decisione e dare la prob.tà di errore di quello migliore nel caso che $p(a_1) = 0{,}1$

Tutti i possibili schemi sono:

$b_1 \to a_1$, $b_2 \to a_1$

$b_1 \to a_1$, $b_2 \to a_2$

$b_1 \to a_2$, $b_2 \to a_2$

$b_1 \to a_2$, $b_2 \to a_1$

Applico lo schema dell'osservatore ideale:

Ricavo $E|U$:

$$E|U = \begin{pmatrix} \frac{0{,}1 \cdot 0{,}7}{p(b_1)} & \frac{0{,}1 \cdot 0{,}3}{p(b_2)} \\ \frac{0{,}9 \cdot 0{,}3}{p(b_1)} & \frac{0{,}9 \cdot 0{,}7}{p(b_2)} \end{pmatrix}$$

Lo schema migliore è: $b_1 \to a_2$, $b_2 \to a_2$

poiché lo schema è degenere $p_{en} = p(a_1) = 0{,}1$

2008 - 5

3) Dato il canale

	b_1	b_2
a_1	0,9	0,1
a_2	0,8	0,2

Enunciare tutti i possibili schemi di decisione e dare la prob.tà di errore di quello migliore nel caso che a_1 e a_2 siano equiprobabili.

Schemi di decisione:

$b_1 \to a_1$, $b_2 \to a_2$

$b_1 \to a_1$, $b_2 \to a_1$

$b_1 \to a_2$, $b_2 \to a_1$

$b_1 \to a_2$, $b_2 \to a_2$

Applico lo schema dell'osservatore ideale:

$$E|U = \begin{pmatrix} \frac{0{,}5 \cdot 0{,}9}{p(b_1)} & \frac{0{,}5 \cdot 0{,}1}{p(b_2)} \\ \frac{0{,}5 \cdot 0{,}8}{p(b_1)} & \frac{0{,}5 \cdot 0{,}2}{p(b_2)} \end{pmatrix}$$

lo schema migliore è $b_1 \to a_1$, $b_2 \to a_2$

$p_{er} = 0{,}5 \cdot 0{,}8 + 0{,}5 \cdot 0{,}1 = 0{,}40 + 0{,}05 = 0{,}45$

2008-6

3) Dare un esempio di canale con capacità 256

In un canale lossless $C = \log r$ dove r è il n° di simboli in entrata, perciò in questo caso $r = 2^{256}$ e il canale avrà un solo elemento $\neq \emptyset$ per ogni colonna.

$$
\begin{array}{c|cccccc}
 & b_1 & b_2 & \cdots & b_{2^{256}-1} & b_{2^{256}} & b_{2^{256}+1} \\
\hline
a_1 & 1 & & & & & \\
a_2 & & 1 & & \emptyset & & \\
\vdots & & & & & & \\
 & & \emptyset & & & & \\
a_{2^{256}-1} & & & & 1 & & \\
a_{2^{256}} & & & & & 1/2 & 1/2
\end{array}
$$

2008-2

4) Dato il canale

$$
\begin{array}{cc|cc}
 & & b_1 & b_2 \\
0.3 & a_1 & 0.3 & 0.7 \\
0.7 & a_2 & 0.6 & 0.4
\end{array}
$$

enunciare tutti i poss. schemi di decisione e dare la prob. tà di errore di quello migliore nel caso $p(a_1) = 0.3$

Schemi di decisione:
$b_1 \to a_1$, $b_2 \to a_1$
$b_1 \to a_1$, $b_2 \to a_2$
$b_1 \to a_2$, $b_2 \to a_2$
$b_1 \to a_2$, $b_2 \to a_1$

$\longrightarrow$

Applico lo schema dell'osservatore ideale:

$$E|U = \begin{pmatrix} \frac{0{,}3 \cdot 0{,}3}{p(b_1)} & \frac{0{,}3 \cdot 0{,}7}{p(b_2)} \\ \frac{0{,}7 \cdot 0{,}6}{p(b_1)} & \frac{0{,}7 \cdot 0{,}4}{p(b_2)} \end{pmatrix}$$

lo schema migliore è: $b_1 \to a_2$, $b_2 \to a_2$

poiché lo schema è deterministico $p_{er} = p(a_1) = 0{,}3$

2009-4

3) Dare un esempio di canale lossless con capacità 2 e uno di canale deterministico con capacità 3.

lossless:

$C = \log_2 r$, r = n° simboli in entrata $\Rightarrow$ $r = 4$, un solo elem. $\neq 0$ per colonna:

	b_1	b_2	b_3	b_4	b_5
a_1	1				
a_2		1		$\emptyset$	
a_3	$\emptyset$		1		
a_4				1/2	1/2

Deterministico:

$C = \log_2 s'$, s' = n° simboli in uscita utilizzati $\Rightarrow$ $s' = 8$, un solo elemento $\neq 0$ per riga.

	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
a_1	0	1	0	0	0	0	0	0	0
a_2	1	0	0	0	0	0	0	0	0
⋮	0	0	0	1	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	0	0	0	0	1	0	0	0	0
a_8	0	0	0	0	0	0	0	0	1

2007-0

CODICI SISTEMATICI E PERFETTI

4) Dare un esempio di codice non perfetto

0000
1111

non è perfetto poiché $n = 4$ (pari)

i codici perfetti con 2 parole devono avere come caratteristiche: $e = m$

$d = 2m + 1$

$n = d$

$t = 2m$

2007-1 / 2008-6 4)

4) Dare un esempio di codice perfetto

000
111

è un codice perfetto del primo tipo che soddisfa il limite di Hamming come uguale:

2 parole

$e = m$

$d = 2m + 1$

$n = d$

$t = 2m$

2007-4 / 2008-4 4)

4) Dare un esempio di codice e CCP sistematico e perfetto.

000
111

è un codice perfetto del primo tipo:

2 parole, $e = m$, $d = 2m + 1$, $n = d$, $t = 2m$

è sistematico di conseguenza poiché A è:

$$A = \left(\begin{array}{c|cc} 1 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right)$$

I ha rango t

2008-5 / 2009-4 4)

4) Dare un esempio di codice perfetto non sistematico.

I codici perfetti di II tipo sono tali che la matrice A ha tutte le possibili colonne di t bit tranne la nulla $\Rightarrow d=3$

$$A = \left(\begin{array}{cccc|ccc} 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \end{array}\right) \qquad t=3 \qquad n = 2^t - 1 = 7$$

matrice degenere $\Rightarrow$ NON sistematico

QUADERNO

Definire un codice sistematico di lunghezza 7 che corregge 1 errore ; con t=3

$$A = \left(\begin{array}{cccc|ccc} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

non degenere

$d = 3 = 2 \cdot 1 + 1$

$n = 7$

$t = 3$

2007-0 **CCP su GF[q] + CODICI CICLICI**

5) Trovare un CCP più piccolo su GF[3] che contiene le seguenti parole: 110, 202, 011. Dire se il codice ottenuto è ciclico.

110
202
011

sono linearmente indipendenti, perciò la matrice G è così fatta:

$$G = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

Per controllare se il codice è ciclico dovrei calcolare tutte le parole, che sono $3^3 = 27$ (tutte le possibili comb. lineari delle righe di G), se manca qualche parola il codice non è ciclico.

2007-1 / 2008-5 5)

5) Trovare il CCP più piccolo su GF[2] che contiene le seguenti parole: 110011, 111100, 001111. Dire se il codice ottenuto è ciclico.

Le tre parole sono linearmente dipendenti:

$$\begin{array}{r} 110011 \\ 001111 \\ \hline 111100 \end{array}$$

quindi G è così fatta: $G = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$

Per verificare la ciclicità osserviamo tutte le possibili parole:

000000
110011
001111
111100

sono $2^k = 2^2 = 4$

non è ciclico poiché per esempio manca 011110

2007-2 / 2008-1 4)

4) Dato il codice a controllo di parità definito dalla matrice

$$A = \left(\begin{array}{cc|ccc} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{array}\right)$$

Determinare la distanza minima e la prob. tà di errore dello schema della max verosimiglianza per un BSC con $p = 99$, $q = 0{,}1$

$d \neq 1$ - non c'è la colonna nulla

$d \neq 2$ - non ci sono due colonne uguali

$d = 3$ - 3 è la cardinalità del più piccolo numero di colonne linearmente dipendenti ($1^a, 2^a, 3^a$).
Il codice corregge almeno 1 errore.

Poiché $t = 3$, tutte le possibili ricezioni sono $2^3 = 8$

la probabilità di errore è:

$$p_{err} = 1 - \Big(1 \cdot p^5 + \underset{\substack{\shortparallel \\ 5}}{\binom{5}{1}} p^4 q + \underset{\substack{\downarrow \\ \text{parole di peso 2 in coda}}}{2} p^3 q^2\Big) = 1 - (0{,}93312) = 0{,}06688$$

2007-3 / 2008-3 5)

5) Dare un esempio di codice ciclico in GF[4]

000
111
ααα
$\alpha^2\alpha^2\alpha^2$

la matrice G associata al codice è:

$G = (1 | 1 1)$

che equivale a scegliere $g(x) = x^2 + x + 1$

Il codice è ciclico perché per qualunque shift le parole restano invariate e sono sempre parole di codice.

2007 - 5

4) Trovare la velocità di trasmissione e la prob.tà di errore di un CCP formato dalle parole 000000 111111, usato per trasmettere due simboli equiprobabili su un canale BSC, con $p=0{,}9$ e $q=0{,}1$

$R = \frac{K}{n} = \frac{1}{6}$ $\qquad G=(1|11111)$

$d=6$, il codice rileva 3 errori e ne corregge almeno 2.

Tutte le possibili sindromi sono $2^5 = 32$, quindi la prob.tà di errore è:

$$p_{err} = 1 - \Big(p^6 + \underbrace{\binom{6}{1}}_{=6} p^5 q + \underbrace{\binom{6}{2}}_{=15} p^4 q^2 + 10 p^3 q^3\Big) =$$

$$= 1 - (0{,}9^6 + 6 \cdot 0{,}9^5 \cdot 0{,}1 + 15 \cdot 0{,}9^4 \cdot 0{,}1^2 + 10 \cdot 0{,}9^3 \cdot 0{,}1^3) =$$

$$= 0{,}00856$$

$$\binom{6}{2} = \frac{6 \cdot 5 \cdot \cancel{4 \cdot 3 \cdot 2}}{2 \cdot \cancel{4 \cdot 3 \cdot 2}} \quad 15$$

$$\binom{6}{3} = \frac{6 \cdot 5 \cdot 4 \cdot \cancel{3 \cdot 2}}{\cancel{3 \cdot 2} \cdot \cancel{3 \cdot 2}} = 20$$

2007 - 5 | 2008 - 0 5)

5) Dare un esempio di codice ciclico che corregge almeno 4 errori.

000000000
111111111

$d = 9 = 2 \cdot 4 + 1$

$G = (1 | 11111111)$

il codice è ciclico perché per ogni shift le parole rimangono le stesse e quindi appartengono al codice.

2007-6

4) Dare un esempio di codice a controllo di parità con velocità di trasmissione 1/5.

00000
11111

con matrice G associata:

$G = (1 \mid 1\,1\,1\,1)$

$R = \frac{k}{n} = \frac{1}{5}$

2007-6 | 2008-2 5) / 2008-2 5)

5) Sapendo che $x^3 - 1 = (x+2)^3$ modulo 3, dare un esempio di codice ciclico non banale di lunghezza 3 su GF[3]

000
111
222

con matrice G associata:

$G = (1 \mid 1\,1)$

equivale a scegliere $g(x) = x^2 + x + 1$

Il codice è ciclico perché per qualsiasi shift circolare le parole restano le stesse e quindi $\in C$.

$g(x) = (x+2) \cdot (x+2) = x^2 + 4x + 4 = x^2 + x + 1$

2008 - 0

4) Trovare la velocità di trasmissione e la prob.tà di errore di un CCP formato dalle parole 00000 11111, usato per trasmettere due simboli equiprobabili su un canale BSC, con $p = 0{,}9$ e $q = 0{,}1$.

$$R = \frac{K}{n} = \frac{1}{5}$$

Poiché $d = 5$ il codice corregge 2 errori

Tutte le possibili sindromi sono $2^4 = 16$

$$\Rightarrow p_{err} = 1 - \Big(\underset{\underset{1}{\parallel}}{p^5} + \underset{\underset{5}{\parallel}}{5p^4q} + \underset{\underset{10}{\parallel}}{\tbinom{5}{2}} p^3 q^2 \Big) = 0{,}00856$$

$$1 + 5 + 10 = 16$$

$$\binom{5}{2} = \frac{5 \cdot 4 \cdot \cancel{3 \cdot 2}}{2 \cdot \cancel{3 \cdot 2}}$$

2009 - 3

5) Trovare un codice ciclico su GF[8] che contenga almeno 8 parole, una qualunque definizione non ambigua del codice è accettabile.

$$G = (1 \mid 11) \qquad C = \{000, 111, \alpha\alpha\alpha, \alpha^2\alpha^2\alpha^2, \alpha^3\alpha^3\alpha^3, \alpha^4\alpha^4\alpha^4, \alpha^5\alpha^5\alpha^5, \alpha^6\alpha^6\alpha^6\}$$

il codice è ciclico poiché per qualunque shift circolare le parole restano le stesse e quindi $\in C$.