

Informazione, entropia e comprimibilità

Michela Barsanti e Luca Sbano
(Licei *Vittoria Colonna*)

Roma 13 Marzo 2020

Indice

Contesto e motivazioni

Aleatorio e Casuale

Modalità di lavoro

Il gioco del 63

Una variante del gioco del 63

Contenuto di informazione secondo Shannon

Giochiamo con l'informazione

Caccia al sottomarino

L'entropia

Osservazioni finali

Bibliografia

Contesto

- Una Classe 4^o di un liceo scientifico delle scienze applicate.
- La classe ha già svolto un modulo di probabilità e statistica bivariata, conosce la funzione logaritmo.
- L'attività ha avuto una durata complessiva di *****.

Motivazioni

- Esplorare il significato di cosa sia l'*informazione* e della sua possibile misura.
- La nozione di entropia in vari contesti...
- Porre il problema della comprimibilità dell'informazione...

Aleatorio e Casuale

Aleatorio e casuale possono essere utilizzati come sinonimi in un contesto matematico ma occorre tener presente alcune sottigliezze...

Per cominciare (*Vocabolario Treccani*)

aleatorio agg. [dal lat. aleatorius, der. di alea gioco di dadi]. Rischioso, incerto: [...] Nel calcolo delle probabilità è sinon. di casuale o **stocastico**: variabile o numero a.; processo a. o stocastico, modello matematico impiegato per studiare l'evolversi nel tempo dei fenomeni dipendenti da fattori casuali, per es. nei problemi delle assicurazioni, nella descrizione del comportamento delle particelle, ecc.

casuale agg. [dal lat. tardo casualis, der. di casus -us caso]. Che dipende dal caso, che avviene o si fa per caso: un incontro c.; per una c. sbadataggine; [...] In statistica, nel calcolo delle probabilità, nelle applicazioni tecniche e nelle scienze sperimentali, che ha il carattere dell'imprevedibilità, che non è descrivibile in modo determinato ma obbedisce a leggi statistiche probabilistiche: fenomeno c. (o aleatorio); processo c. (o aleatorio), modello matematico di un fenomeno casuale costruito mediante gli strumenti della probabilità e della statistica;...

Un punto interessante...

- **stocastico** agg. [dal gr. *στοχαικός* congetturale, propr. che mira bene, abile nel congetturare, der. di *στοχαζομαί* mirare, congetturare da *στοχός* bersaglio, mira, congettura]

L'aggettivo **stocastico** rimanda ad una doppio aspetto uno deterministico ed uno non deterministico sempre presenti nei fenomeni...

Un punto interessante...

- **stocastico** agg. [dal gr. *στοχαικός* congetturale, propr. che mira bene, abile nel congetturare, der. di *στοχαζομαί* mirare, congetturare da *στοχός* bersaglio, mira, congettura]

L'aggettivo **stocastico** rimanda ad una doppio aspetto uno deterministico ed uno non deterministico sempre presenti nei fenomeni...

Notazioni

- Chiamiamo **stringa** un insieme finito di caratteri o numeri.
- Chiamiamo **successione** un insieme di caratteri che non hanno un elemento ultimo (*stringe di infiniti caratteri*).

Se si utilizza il solo termine **successione** è importante che gli studenti osservino la distinzione fra *successioni* e *successioni finite*.

In generale si nota che la capacità degli studenti di cogliere questo aspetto non appare in modo naturale soprattutto nelle classi iniziali.

Notazioni

- Chiamiamo **stringa** un insieme finito di caratteri o numeri.
- Chiamiamo **successione** un insieme di caratteri che non hanno un elemento ultimo (*stringe di infiniti caratteri*).

Se si utilizza il solo termine **successione** è importante che gli studenti osservino la distinzione fra *successioni* e *successioni finite*.

In generale si nota che la capacità degli studenti di cogliere questo aspetto non appare in modo naturale soprattutto nelle classi iniziali.

Modalità di lavoro

- In classe si possono far lavorare gli studenti in gruppi da tre.
- Si propongono tre giochi.

Modalità di lavoro

- In classe si possono far lavorare gli studenti in gruppi da tre.
- Si propongono tre giochi.

Scheda di lavoro

Ai gruppi viene proposto il seguente gioco: *Uno di voi immagini di scegliere un numero fra 0 e 63 (estremi inclusi). Gli altri studenti devono cercare di indovinare il numero ponendo delle semplici domande a cui si possa rispondere con un si o con un no senza ulteriori commenti.*

- 1 Qual è il numero massimo di domande si possono fare per ottenere la risposta?
- 2 Qual è il numero minimo di domande che si possono fare per indovinare?
- 3 Quale relazione si può immaginare fra il numero di elementi dell'insieme $A = \{0, \dots, 63\}$ e il numero minimo (che indicheremo con $I(A)$) di domande trovate?

Prima parte

- Nei gruppi a rotazione uno pensa un numero nel range assegnato $[n, n + 63]$, gli altri due porranno una domanda a turno.
- 10 minuti per indovinare tutti e tre i giocatori del gruppo.
- 10 minuti: discussione regolamentata.
- Il docente chiede chi ha ottimizzato la strategia.
- Il docente proietta una slide con:
Ho pensato un numero, per fare prima nell'intervallo di ampiezza metà.
- Gli alunni devono stimare il numero di domande e poi cominciare a porle una per banco in modo da verificare la compressione del processo.

Seconda parte

- 15 minuti scheda con le due tabelle 41, 20 con commento e conclusioni.
- 15 minuti per $A \times B$ esempio e generalizzazione
- Domanda da lasciare aperta: Quando un prodotto diventa una somma? Conosci una funzione che si comporta così?

Strategia 1

- La richiesta di domande binarie (si no) suggerisce di usare le relazioni d'ordine \geq, \leq e cercare di avvicinarsi al numero nascosto con un procedimento per eccesso e per difetto.
- Si può per esempio dividere il numero di elementi in A per 2 e chiedere se $x \geq |A|/2 = 32$. Poi se la risposta è positiva vorrà dire che $x \in [32, 63]$, la risposta negativa implicherà $x \in [0, 32]$.
- La procedura potrà essere iterata.
- qualè il numero minimo di iterazioni $I(A)$ per trovare il numero?

Strategia 1

- La richiesta di domande binarie (si no) suggerisce di usare le relazioni d'ordine \geq, \leq e cercare di avvicinarsi al numero nascosto con un procedimento per eccesso e per difetto.
- Si può per esempio dividere il numero di elementi in A per 2 e chiedere se $x \geq |A|/2 = 32$. Poi se la risposta è positiva vorrà dire che $x \in [32, 63]$, la risposta negativa implicherà $x \in [0, 32]$.
- La procedura potrà essere iterata.
- **qualè il numero minimo di iterazioni $I(A)$ per trovare il numero?**

Strategia 2

Posto $x \in \{0, \dots, 63\}$, ecco un'altra possibile strategia:

① $x \geq 32 \pmod{64}$

② $x \geq 16 \pmod{32}$

③ $x \geq 8 \pmod{16}$

④ $x \geq 4 \pmod{8}$

⑤ $x \geq 2 \pmod{4}$

⑥ $x \geq 1 \pmod{2}$

Il numero di domande $I(A)$ binarie è 6

Qual è la relazione fra il numero di elementi in A ed il numero di domande (si,no)?

Se ad ogni risposta positiva si assegna 1 e 0 per quelle negative, a cosa corrisponde la sequenza di 1 e 0 così ottenuta?

Il numero di domande $I(A)$ binarie è 6

Qual è la relazione fra il numero di elementi in A ed il numero di domande (si,no)?

Se ad ogni risposta positiva si assegna 1 e 0 per quelle negative, a cosa corrisponde la sequenza di 1 e 0 così ottenuta?

Il numero di domande $I(A)$ binarie è 6

Qual è la relazione fra il numero di elementi in A ed il numero di domande (si,no)?

Se ad ogni risposta positiva si assegna 1 e 0 per quelle negative, a cosa corrisponde la sequenza di 1 e 0 così ottenuta?

Il numero di domande $I(A)$ binarie è 6

Qual è la relazione fra il numero di elementi in A ed il numero di domande (si,no)?

Se ad ogni risposta positiva si assegna 1 e 0 per quelle negative, a cosa corrisponde la sequenza di 1 e 0 così ottenuta?

41

Ho pensato 41	Modulo	Risposta	Si= 1, No= 0	Valore
$x \geq 32$	mod 64			
$x \geq 16$	mod 32			
$x \geq 8$	mod 16			
$x \geq 4$	mod 8			
$x \geq 2$	mod 4			
$x \geq 1$	mod 2			

- Quale informazione puoi trarre da ogni riga?
- Qual informazione complessiva puoi indovinare?

20

Ho pensato 20	Modulo	Risposta	Si= 1, No= 0	Valore
$x \geq 32$	mod 64			
$x \geq 16$	mod 32			
$x \geq 8$	mod 16			
$x \geq 4$	mod 8			
$x \geq 2$	mod 4			
$x \geq 1$	mod 2			

- Quale informazione puoi trarre da ogni riga?
- Qual informazione complessiva puoi indovinare?

Una variante del gioco del 63

- Si prendano per esempio due insiemi $B = \{0, 1, 2, 3, 5, 6, 7\}$ e $C = \{0, 1, 2, 3\}$ si può lo stesso tipo di gioco sia con B che con C e determinare $I(B)$ e $I(C)$.
- Si può poi costruire l'insieme prodotto $B \times C$ i cui elementi possono essere messi in corrispondenza biunivoca con gli interi da 0 a 36.
- Scelto un elemento di $B \times C$, in quante domande lo si può determinare?
- Che relazione si può immaginare fra $I(B)$, $I(C)$ e $I(B \times C)$?
- Quale funzione conosci che soddisfa una relazione del tipo:

$$I(B \times C) = I(B) + I(C) \quad ?(\text{Fine prima lezione})$$

Una variante del gioco del 63

- Si prendano per esempio due insiemi $B = \{0, 1, 2, 3, 5, 6, 7\}$ e $C = \{0, 1, 2, 3\}$ si può lo stesso tipo di gioco sia con B che con C e determinare $I(B)$ e $I(C)$.
- Si può poi costruire l'insieme prodotto $B \times C$ i cui elementi possono essere messi in corrispondenza biunivoca con gli interi da 0 a 36.
- **Scelto un elemento di $B \times C$, in quante domande lo si può determinare?**
- Che relazione si può immaginare fra $I(B)$, $I(C)$ e $I(B \times C)$?
- Quale funzione conosci che soddisfa una relazione del tipo:

$$I(B \times C) = I(B) + I(C) \quad ?(\text{Fine prima lezione})$$

Una variante del gioco del 63

- Si prendano per esempio due insiemi $B = \{0, 1, 2, 3, 5, 6, 7\}$ e $C = \{0, 1, 2, 3\}$ si può lo stesso tipo di gioco sia con B che con C e determinare $I(B)$ e $I(C)$.
- Si può poi costruire l'insieme prodotto $B \times C$ i cui elementi possono essere messi in corrispondenza biunivoca con gli interi da 0 a 36.
- Scelto un elemento di $B \times C$, in quante domande lo si può determinare?
- Che relazione si può immaginare fra $I(B)$, $I(C)$ e $I(B \times C)$?
- Quale funzione conosci che soddisfa una relazione del tipo:

$$I(B \times C) = I(B) + I(C) \quad ?(\text{Fine prima lezione})$$

Una variante del gioco del 63

- Si prendano per esempio due insiemi $B = \{0, 1, 2, 3, 5, 6, 7\}$ e $C = \{0, 1, 2, 3\}$ si può lo stesso tipo di gioco sia con B che con C e determinare $I(B)$ e $I(C)$.
- Si può poi costruire l'insieme prodotto $B \times C$ i cui elementi possono essere messi in corrispondenza biunivoca con gli interi da 0 a 36.
- Scelto un elemento di $B \times C$, in quante domande lo si può determinare?
- Che relazione si può immaginare fra $I(B)$, $I(C)$ e $I(B \times C)$?
- Quale funzione conosci che soddisfa una relazione del tipo:

$$I(B \times C) = I(B) + I(C) \quad ?(\text{Fine prima lezione})$$

Contenuto di informazione secondo Shannon

- La funzione $I(x)$ misura la quantità d'informazione contenuta in $x \in \mathcal{A}$.
- La funzione $I(x)$ soddisfa le proprietà della funzione logaritmo.
- Shannon definisce il *Contenuto di informazione* come:

$$I(x) = \log_2 \left(\frac{1}{p(x)} \right).$$

dove $p(x)$ è la probabilità che un elemento di \mathcal{A} sia scelto a caso. Se le x sono equiprobabili, allora $p(x) = 1/|\mathcal{A}|$ e $I(x) = \log_2(|\mathcal{A}|)$.

- Notare che il logaritmo rende l'*informazione* $I(\cdot)$ additiva.

Contenuto di informazione secondo Shannon

- La funzione $I(x)$ misura la quantità d'informazione contenuta in $x \in \mathcal{A}$.
- La funzione $I(x)$ soddisfa le proprietà della funzione logaritmo.
- Shannon definisce il *Contenuto di informazione* come:

$$I(x) = \log_2 \left(\frac{1}{p(x)} \right).$$

dove $p(x)$ è la probabilità che un elemento di \mathcal{A} sia scelto a caso. Se le x sono equiprobabili, allora $p(x) = 1/|\mathcal{A}|$ e $I(x) = \log_2(|\mathcal{A}|)$.

- Notare che il logaritmo rende l'*informazione* $I(\cdot)$ additiva.

Contenuto di informazione secondo Shannon

- La funzione $I(x)$ misura la quantità d'informazione contenuta in $x \in \mathcal{A}$.
- La funzione $I(x)$ soddisfa le proprietà della funzione logaritmo.
- Shannon definisce il *Contenuto di informazione* come:

$$I(x) = \log_2 \left(\frac{1}{p(x)} \right).$$

dove $p(x)$ è la probabilità che un elemento di \mathcal{A} sia scelto a caso. Se le x sono equiprobabili, allora $p(x) = 1/|\mathcal{A}|$ e $I(x) = \log_2(|\mathcal{A}|)$.

- Notare che il logaritmo rende l'*informazione* $I(\cdot)$ additiva.

Contenuto di informazione secondo Shannon

- La funzione $I(x)$ misura la quantità d'informazione contenuta in $x \in \mathcal{A}$.
- La funzione $I(x)$ soddisfa le proprietà della funzione logaritmo.
- Shannon definisce il *Contenuto di informazione* come:

$$I(x) = \log_2 \left(\frac{1}{p(x)} \right).$$

dove $p(x)$ è la probabilità che un elemento di \mathcal{A} sia scelto a caso. Se le x sono equiprobabili, allora $p(x) = 1/|\mathcal{A}|$ e $I(x) = \log_2(|\mathcal{A}|)$.

- Notare che il logaritmo rende l'*informazione* $I(\cdot)$ additiva.

Contenuto di informazione secondo Shannon

- La funzione $I(x)$ misura la quantità d'informazione contenuta in $x \in \mathcal{A}$.
- La funzione $I(x)$ soddisfa le proprietà della funzione logaritmo.
- Shannon definisce il *Contenuto di informazione* come:

$$I(x) = \log_2 \left(\frac{1}{p(x)} \right).$$

dove $p(x)$ è la probabilità che un elemento di \mathcal{A} sia scelto a caso. Se le x sono equiprobabili, allora $p(x) = 1/|\mathcal{A}|$ e $I(x) = \log_2(|\mathcal{A}|)$.

- Notare che il logaritmo rende l'*informazione* $I(\cdot)$ additiva.

Esempi

- Moneta non truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/2$, $p(C) = 1/2$, quanto quanto vale l'informazione?
- Moneta truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/3$, $p(C) = 2/3$, quanto quanto vale l'informazione?
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate da coppie **distinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate da coppie **indistinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate doppiette e poi triplete...

Esempi

- Moneta non truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/2$, $p(C) = 1/2$, quanto quanto vale l'informazione?
- Moneta truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/3$, $p(C) = 2/3$, quanto quanto vale l'informazione?
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate da coppie **distinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate da coppie **indistinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate doppiette e poi triplete...

Esempi

- Moneta non truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/2$, $p(C) = 1/2$, quanto quanto vale l'informazione?
- Moneta truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/3$, $p(C) = 2/3$, quanto quanto vale l'informazione?
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate da coppie **distinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate da coppie **indistinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate doppiette e poi triplete...

Esempi

- Moneta non truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/2$, $p(C) = 1/2$, quanto quanto vale l'informazione?
- Moneta truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/3$, $p(C) = 2/3$, quanto quanto vale l'informazione?
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate da coppie **distinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate da coppie **indistinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate doppiette e poi triplete...

Esempi

- Moneta non truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/2$, $p(C) = 1/2$, quanto quanto vale l'informazione?
- Moneta truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/3$, $p(C) = 2/3$, quanto quanto vale l'informazione?
- Stringhe di elemnti di $\mathcal{A} = \{0, 1\}$ formate da coppie **distinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elemnti di $\mathcal{A} = \{0, 1\}$ formate da coppie **indistinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate doppiette e poi triplete...

Esempi

- Moneta non truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/2$, $p(C) = 1/2$, quanto quanto vale l'informazione?
- Moneta truccata: $\mathcal{A} = \{T, C\}$ con $p(T) = 1/3$, $p(C) = 2/3$, quanto quanto vale l'informazione?
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate da coppie **distinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate da coppie **indistinguibili**. Con quali probabilità? (Considerate per esempio $p(0) = p(1) = 1/2$).
- Stringhe di elementi di $\mathcal{A} = \{0, 1\}$ formate doppiette e poi triplete...

Lettere e microstati, stringhe e macrostati

- \mathcal{A} è un insieme di lettere. Un insieme (finito) di lettere forma una stringa.
- È utile introdurre una funzione **stato** che sceglie un elemento in \mathcal{A} . Le possibili lettere di un alfabeto costituiscono i possibili **microstati**.
- Una stringa è formata da tante lettere quindi è formata da tanti microstati.
- Considero l'insieme \mathcal{A}^L di tutte le stringhe di lunghezza L , questo insieme può essere partizionato ed ogni partizione sarà formata da gruppi di stringhe diverse, ogni elemento della partizione è chiamato **macrostato**.

Esempio (1)

- $\mathcal{A} = \{T, C\}$, i microstati sono T e C .
- Quali sono le possibili stringhe di lunghezza 2?
- $\mathcal{A}^2 = \{TT, TC, CT, CC\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^2 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	TT	TC	CT	CC
Probabilità	1/4	1/4	1/4	1/4

- Oppure i macrostati:

Macrostati	TT	{TC,CT}	CC
Probabilità	1/4	2/4	1/4

Esempio (1)

- $\mathcal{A} = \{T, C\}$, i microstati sono T e C .
- Quali sono le possibili stringhe di lunghezza 2?
- $\mathcal{A}^2 = \{TT, TC, CT, CC\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^2 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	TT	TC	CT	CC
Probabilità	1/4	1/4	1/4	1/4

- Oppure i macrostati:

Macrostati	TT	{TC,CT}	CC
Probabilità	1/4	2/4	1/4

Esempio (1)

- $\mathcal{A} = \{T, C\}$, i microstati sono T e C .
- Quali sono le possibili stringhe di lunghezza 2?
- $\mathcal{A}^2 = \{TT, TC, CT, CC\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^2 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	TT	TC	CT	CC
Probabilità	1/4	1/4	1/4	1/4

- Oppure i macrostati:

Macrostati	TT	{TC,CT}	CC
Probabilità	1/4	2/4	1/4

Esempio (1)

- $\mathcal{A} = \{T, C\}$, i microstati sono T e C .
- Quali sono le possibili stringhe di lunghezza 2?
- $\mathcal{A}^2 = \{TT, TC, CT, CC\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^2 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	TT	TC	CT	CC
Probabilità	1/4	1/4	1/4	1/4

- Oppure i macrostati:

Macrostati	TT	{TC,CT}	CC
Probabilità	1/4	2/4	1/4

Esempio (1)

- $\mathcal{A} = \{T, C\}$, i microstati sono T e C .
- Quali sono le possibili stringhe di lunghezza 2?
- $\mathcal{A}^2 = \{TT, TC, CT, CC\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^2 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	TT	TC	CT	CC
Probabilità	1/4	1/4	1/4	1/4

- Oppure i macrostati:

Macrostati	TT	{TC,CT}	CC
Probabilità	1/4	2/4	1/4

Esempio (1)

- $\mathcal{A} = \{T, C\}$, i microstati sono T e C .
- Quali sono le possibili stringhe di lunghezza 2?
- $\mathcal{A}^2 = \{TT, TC, CT, CC\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^2 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	TT	TC	CT	CC
Probabilità	1/4	1/4	1/4	1/4

- Oppure i macrostati:

Macrostati	TT	{TC,CT}	CC
Probabilità	1/4	2/4	1/4

Esempio (1)

- $\mathcal{A} = \{T, C\}$, i microstati sono T e C .
- Quali sono le possibili stringhe di lunghezza 2?
- $\mathcal{A}^2 = \{TT, TC, CT, CC\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^2 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	TT	TC	CT	CC
Probabilità	1/4	1/4	1/4	1/4

- Oppure i macrostati:

Macrostati	TT	{TC,CT}	CC
Probabilità	1/4	2/4	1/4

Esempio (2)

- $\mathcal{A} = \{1, 0\}$, i microstati sono 1 e 0.
- Quali sono le possibili stringhe di lunghezza 3?
- $\mathcal{A}^3 = \{000, 100, 010, 001, 110, 101, 011, 111\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^3 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	000	{100,010,001}	{110,011,101}	111
Probabilità	1/8	3/8	3/8	1/8

Esempio (2)

- $\mathcal{A} = \{1, 0\}$, i microstati sono 1 e 0.
- Quali sono le possibili stringhe di lunghezza 3?
- $\mathcal{A}^3 = \{000, 100, 010, 001, 110, 101, 011, 111\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^3 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	000	{100,010,001}	{110,011,101}	111
Probabilità	1/8	3/8	3/8	1/8

Esempio (2)

- $\mathcal{A} = \{1, 0\}$, i microstati sono 1 e 0.
- Quali sono le possibili stringhe di lunghezza 3?
- $\mathcal{A}^3 = \{000, 100, 010, 001, 110, 101, 011, 111\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^3 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	000	{100,010,001}	{110,011,101}	111
Probabilità	1/8	3/8	3/8	1/8

Esempio (2)

- $\mathcal{A} = \{1, 0\}$, i microstati sono 1 e 0.
- Quali sono le possibili stringhe di lunghezza 3?
- $\mathcal{A}^3 = \{000, 100, 010, 001, 110, 101, 011, 111\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^3 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	000	{100,010,001}	{110,011,101}	111
Probabilità	1/8	3/8	3/8	1/8

Esempio (2)

- $\mathcal{A} = \{1, 0\}$, i microstati sono 1 e 0.
- Quali sono le possibili stringhe di lunghezza 3?
- $\mathcal{A}^3 = \{000, 100, 010, 001, 110, 101, 011, 111\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^3 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	000	{100,010,001}	{110,011,101}	111
Probabilità	1/8	3/8	3/8	1/8

Esempio (2)

- $\mathcal{A} = \{1, 0\}$, i microstati sono 1 e 0.
- Quali sono le possibili stringhe di lunghezza 3?
- $\mathcal{A}^3 = \{000, 100, 010, 001, 110, 101, 011, 111\}$
- In quanti partizioni possiamo suddividere \mathcal{A}^3 ?
- Una partizione potrebbe contenere i macrostati:

Macrostati	000	{100,010,001}	{110,011,101}	111
Probabilità	1/8	3/8	3/8	1/8

Microstati e macrostati in fisica

- \mathcal{A} è un insieme degli atomi in una stanza. Un tale sistema avrebbe uno **stato** definito in meccanica classica specificando per ogni atomo la sua posizione e velocità. Questo è un **microstato**. Pensate in una mole ci sono circa $N_A \simeq 10^{23}$ particelle!!!
- In effetti in fisica, si mostra che le proprietà di N_A particelle sono definite conoscendo solo sottoinsiemi formati da moltissimi microstati. Gli elementi di una partizione dell'insieme dei microstati è detto **macrostato**.
- Se si fissa l'energia E di un sistema, si fissano anche tutti i possibili microstati di un sistema compatibili con E .
- Un **macrostato** è invece definito dalle grandezze macroscopiche come P , V e T , ed è formato dall'insieme di tutti i possibili microstati compatibili.

Microstati e macrostati in fisica

- \mathcal{A} è un insieme degli atomi in una stanza. Un tale sistema avrebbe uno **stato** definito in meccanica classica specificando per ogni atomo la sua posizione e velocità. Questo è un **microstato**. Pensate in una mole ci sono circa $N_A \simeq 10^{23}$ particelle!!!
- In effetti in fisica, si mostra che le proprietà di N_A particelle sono definite conoscendo solo sottoinsiemi formati da moltissimi microstati. Gli elementi di una partizione dell'insieme dei microstati è detto **macrostato**.
- Se si fissa l'energia E di un sistema, si fissano anche tutti i possibili microstati di un sistema compatibili con E .
- Un **macrostato** è invece definito dalle grandezze macroscopiche come P , V e T , ed è formato dall'insieme di tutti i possibili microstati compatibili.

Microstati e macrostati in fisica

- \mathcal{A} è un insieme degli atomi in una stanza. Un tale sistema avrebbe uno **stato** definito in meccanica classica specificando per ogni atomo la sua posizione e velocità. Questo è un **microstato**. Pensate in una mole ci sono circa $N_A \simeq 10^{23}$ particelle!!!
- In effetti in fisica, si mostra che le proprietà di N_A particelle sono definite conoscendo solo sottoinsiemi formati da moltissimi microstati. Gli elementi di una partizione dell'insieme dei microstati è detto **macrostato**.
- Se si fissa l'energia E di un sistema, si fissano anche tutti i possibili microstati di un sistema compatibili con E .
- Un **macrostato** è invece definito dalle grandezze macroscopiche come P , V e T , ed è formato dall'insieme di tutti i possibili microstati compatibili.

Microstati e macrostati in fisica

- \mathcal{A} è un insieme degli atomi in una stanza. Un tale sistema avrebbe uno **stato** definito in meccanica classica specificando per ogni atomo la sua posizione e velocità. Questo è un **microstato**. Pensate in una mole ci sono circa $N_A \simeq 10^{23}$ particelle!!!
- In effetti in fisica, si mostra che le proprietà di N_A particelle sono definite conoscendo solo sottoinsiemi formati da moltissimi microstati. Gli elementi di una partizione dell'insieme dei microstati è detto **macrostato**.
- Se si fissa l'energia E di un sistema, si fissano anche tutti i possibili microstati di un sistema compatibili con E .
- Un **macrostato** è invece definito dalle grandezze macroscopiche come P , V e T , ed è formato dall'insieme di tutti i possibili microstati compatibili.

Microstati e macrostati in fisica

- \mathcal{A} è un insieme degli atomi in una stanza. Un tale sistema avrebbe uno **stato** definito in meccanica classica specificando per ogni atomo la sua posizione e velocità. Questo è un **microstato**. Pensate in una mole ci sono circa $N_A \simeq 10^{23}$ particelle!!!
- In effetti in fisica, si mostra che le proprietà di N_A particelle sono definite conoscendo solo sottoinsiemi formati da moltissimi microstati. Gli elementi di una partizione dell'insieme dei microstati è detto **macrostato**.
- Se si fissa l'energia E di un sistema, si fissano anche tutti i possibili microstati di un sistema compatibili con E .
- Un **macrostato** è invece definito dalle grandezze macroscopiche come P , V e T , ed è formato dall'insieme di tutti i possibili microstati compatibili.

Storia dell'entropia: L. Boltzmann



Figura 1: L. Boltzmann (1844 - 1906)

Dalla meccanica classica alla termodinamica. A fissata energia E tutti i microstati compatibili sono contati da una funzione $\Omega(E)$:

$$S(E) = k_B \ln \Omega(E).$$

Storia dell'entropia: J.W. Gibbs

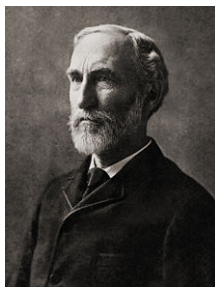


Figura 2: J.W. Gibbs (1839 - 1903)

Generalizza la definizione di entropia. Un macrostato s è subordinata alla realizzazione di k eventi con probabilità p_1, \dots, p_k :

$$H_G(s) = k_B \sum_{i=1}^k p_i \ln \left(\frac{1}{p_i} \right).$$

Storia dell'entropia: C. Shannon



Figura 3: C. Shannon (1916 - 2001)

$x_i \in s$, x_i ha probabilità p_i in s , l'informazione è $I(x_i) = \log_2(1/p_i)$ in *bit*. L'entropia di s

$$H_S(s) = \sum_{i=1}^k p_i \log_2 \left(\frac{1}{p_i} \right).$$

Calcolo dell'entropia di Shannon in \mathcal{A}^2

- $\mathcal{A} = \{T, C\}$, \mathcal{A}_1^2 :

Macrostatì	TT	TC	CT	CC
Probabilità	1/4	1/4	1/4	1/4

- L'informazione:

$$I(TT) = I(TC) = I(CT) = I(CC) = \log_2(4) = 2$$

- L'entropia:

$$H_S(TT) = H_S(TC) = H_S(CT) = H_S(CC) = \frac{1}{4} \log_2(4) = \frac{1}{2}$$

Calcolo dell'entropia di Shannon in \mathcal{A}^2

- $\mathcal{A} = \{T, C\}$, \mathcal{A}_1^2 :

Macrostatì	TT	TC	CT	CC
Probabilità	1/4	1/4	1/4	1/4

- L'informazione:

$$I(TT) = I(TC) = I(CT) = I(CC) = \log_2(4) = 2$$

- L'entropia:

$$H_S(TT) = H_S(TC) = H_S(CT) = H_S(CC) = \frac{1}{4} \log_2(4) = \frac{1}{2}$$

Calcolo dell'entropia di Shannon in \mathcal{A}^2

- $\mathcal{A} = \{T, C\}$, \mathcal{A}_1^2 :

Macrostatì	TT	TC	CT	CC
Probabilità	1/4	1/4	1/4	1/4

- L'informazione:

$$I(TT) = I(TC) = I(CT) = I(CC) = \log_2(4) = 2$$

- L'entropia:

$$H_S(TT) = H_S(TC) = H_S(CT) = H_S(CC) = \frac{1}{4} \log_2(4) = \frac{1}{2}$$

Calcolo dell'entropia di Shannon in \mathcal{A}^2

- $\mathcal{A} = \{T, C\}$, \mathcal{A}_2^2 :

Macrostatì	TT	{TC,CT}	CC
Probabilità	1/4	2/4	1/4

- L'informazione:

$$I(TT) = I(CC) = \log_2(4) = 2, \quad I(\{TC, CT\}) = \log_2(2) = 1$$

- L'entropia:

$$H_S(TT) = H_S(CC) = \frac{1}{4} \log_2(4) = \frac{1}{2}, \quad H_S(\{TC, CT\}) = \frac{1}{2} \log_2(2)$$

Calcolo dell'entropia di Shannon in \mathcal{A}^2

- $\mathcal{A} = \{T, C\}$, \mathcal{A}_2^2 :

Macrostatì	TT	{TC,CT}	CC
Probabilità	1/4	2/4	1/4

- L'informazione:

$$I(TT) = I(CC) = \log_2(4) = 2, \quad I(\{TC, CT\}) = \log_2(2) = 1$$

- L'entropia:

$$H_S(TT) = H_S(CC) = \frac{1}{4} \log_2(4) = \frac{1}{2}, \quad H_S(\{TC, CT\}) = \frac{1}{2} \log_2(2)$$

Calcolo dell'entropia di Shannon in \mathcal{A}^2

- $\mathcal{A} = \{T, C\}$, \mathcal{A}_2^2 :

Macrostatì	TT	{TC,CT}	CC
Probabilità	1/4	2/4	1/4

- L'informazione:

$$I(TT) = I(CC) = \log_2(4) = 2, \quad I(\{TC, CT\}) = \log_2(2) = 1$$

- L'entropia:

$$H_S(TT) = H_S(CC) = \frac{1}{4} \log_2(4) = \frac{1}{2}, \quad H_S(\{TC, CT\}) = \frac{1}{2} \log_2(2)$$

Calcolo dell'entropia di Shannon in \mathcal{A}^3

- $\mathcal{A} = \{1, 0\}$, \mathcal{A}^3 :

Macrostatì	000	{100,010,001}	{110,011,101}	111
Probabilità	1/8	3/8	3/8	1/8

- L'informazione:

$$I(000) = I(111) = \log_2(8) = 3,$$

$$I(\{100, 010, 001\}) = I(\{110, 011, 101\}) = \log_2(8/3)$$

- L'entropia:

$$H_S(TT) = H_S(CC) = \frac{1}{8} \log_2(8) = \frac{3}{8},$$

$$H_S(\{100, 010, 001\}) = H_S(\{110, 011, 101\}) = \frac{3}{8} \log_2(8/3).$$

Calcolo dell'entropia di Shannon in \mathcal{A}^3

- $\mathcal{A} = \{1, 0\}$, \mathcal{A}^3 :

Macrostatì	000	{100,010,001}	{110,011,101}	111
Probabilità	1/8	3/8	3/8	1/8

- L'informazione:

$$I(000) = I(111) = \log_2(8) = 3,$$

$$I(\{100, 010, 001\}) = I(\{110, 011, 101\}) = \log_2(8/3)$$

- L'entropia:

$$H_S(TT) = H_S(CC) = \frac{1}{8} \log_2(8) = \frac{3}{8},$$

$$H_S(\{100, 010, 001\}) = H_S(\{110, 011, 101\}) = \frac{3}{8} \log_2(8/3).$$

Calcolo dell'entropia di Shannon in \mathcal{A}^3

- $\mathcal{A} = \{1, 0\}$, \mathcal{A}^3 :

Macrostatì	000	{100,010,001}	{110,011,101}	111
Probabilità	1/8	3/8	3/8	1/8

- L'informazione:

$$I(000) = I(111) = \log_2(8) = 3,$$

$$I(\{100, 010, 001\}) = I(\{110, 011, 101\}) = \log_2(8/3)$$

- L'entropia:

$$H_S(TT) = H_S(CC) = \frac{1}{8} \log_2(8) = \frac{3}{8},$$

$$H_S(\{100, 010, 001\}) = H_S(\{110, 011, 101\}) = \frac{3}{8} \log_2(8/3).$$

Caccia al sottomarino

La tabella che segue è un classico schema per la *battaglia navale* che contiene un solo sommergibile da colpire, ad esempio posto in (5, 4).

	1	2	3	4	5	6	7	8
1								
2								
3								
4					S			
5								
6								
7								
8								

Caccia al sottomarino

	1	2	3	4	5	6	7	8
1								
2								
3								
4								
5								
6								
7								
8								

Caccia al sottomarino

Quanta informazione è contenuta nello schema della *caccia al sottomarino*?

Caccia al sottomarino

- Qual è la probabilità di colpire al primo colpo il sommergibile?
- Quanta informazione si otterrebbe?
- Si immagini di colpirlo dopo 1 o 2 o 63 colpi mancati come cambierebbe l'informazione ad ogni passo?
- $p_c = 1/64$ e $p_{nc} = 63/64$. Colpire subito significherebbe poter ricostruire tutta la tabella.
- $I = \log_2(64) = 6$ bit
- Ad ogni colpo l'informazione cambia, poichè la probabilità di trovare il sommergibile cambia. Se per esempio si eseguono 33 colpi mancati si ottiene un'informazione pari a:

$$I = \log_2(64/63) + \log_2(63/62) + \dots + \log_2(33/32) = 1 \text{ bit.}$$

Ma se fosse colpito dopo 34 tiri: $I = 1 + \log_2(32/1) = 6$ bit, ritroveremmo tutta l'informazione.

Caccia al sottomarino

- Qual è la probabilità di colpire al primo colpo il sommergibile?
- Quanta informazione si otterrebbe?
- Si immagini di colpirlo dopo 1 o 2 o 63 colpi mancati come cambierebbe l'informazione ad ogni passo?
- $p_c = 1/64$ e $p_{nc} = 63/64$. Colpire subito significherebbe poter ricostruire tutta la tabella.
- $I = \log_2(64) = 6$ bit
- Ad ogni colpo l'informazione cambia, poichè la probabilità di trovare il sommergibile cambia. Se per esempio si eseguono 33 colpi mancati si ottiene un'informazione pari a:

$$I = \log_2(64/63) + \log_2(63/62) + \dots + \log_2(33/32) = 1 \text{ bit.}$$

Ma se fosse colpito dopo 34 tiri: $I = 1 + \log_2(32/1) = 6$ bit, ritroveremmo tutta l'informazione.

Caccia al sottomarino

- Qual è la probabilità di colpire al primo colpo il sommergibile?
- Quanta informazione si otterrebbe?
- Si immagini di colpirlo dopo 1 o 2 o 63 colpi mancati come cambierebbe l'informazione ad ogni passo?
- $p_c = 1/64$ e $p_{nc} = 63/64$. Colpire subito significherebbe poter ricostruire tutta la tabella.
- $I = \log_2(64) = 6$ bit
- Ad ogni colpo l'informazione cambia, poichè la probabilità di trovare il sommergibile cambia. Se per esempio si eseguono 33 colpi mancati si ottiene un'informazione pari a:

$$I = \log_2(64/63) + \log_2(63/62) + \dots + \log_2(33/32) = 1 \text{ bit.}$$

Ma se fosse colpito dopo 34 tiri: $I = 1 + \log_2(32/1) = 6$ bit, ritroveremmo tutta l'informazione.

Caccia al sottomarino

- Qual è la probabilità di colpire al primo colpo il sommergibile?
- Quanta informazione si otterrebbe?
- Si immagini di colpirlo dopo 1 o 2 o 63 colpi mancati come cambierebbe l'informazione ad ogni passo?
- $p_c = 1/64$ e $p_{nc} = 63/64$. Colpire subito significherebbe poter ricostruire tutta la tabella.
- $I = \log_2(64) = 6$ bit
- Ad ogni colpo l'informazione cambia, poichè la probabilità di trovare il sommergibile cambia. Se per esempio si eseguono 33 colpi mancati si ottiene un'informazione pari a:

$$I = \log_2(64/63) + \log_2(63/62) + \dots + \log_2(33/32) = 1 \text{ bit.}$$

Ma se fosse colpito dopo 34 tiri: $I = 1 + \log_2(32/1) = 6$ bit, ritroveremmo tutta l'informazione.

Caccia al sottomarino

- Qual è la probabilità di colpire al primo colpo il sommergibile?
- Quanta informazione si otterrebbe?
- Si immagini di colpirlo dopo 1 o 2 o 63 colpi mancati come cambierebbe l'informazione ad ogni passo?
- $p_c = 1/64$ e $p_{nc} = 63/64$. Colpire subito significherebbe poter ricostruire tutta la tabella.
- $I = \log_2(64) = 6$ bit
- Ad ogni colpo l'informazione cambia, poichè la probabilità di trovare il sommergibile cambia. Se per esempio si eseguono 33 colpi mancati si ottiene un'informazione pari a:

$$I = \log_2(64/63) + \log_2(63/62) + \dots + \log_2(33/32) = 1 \text{ bit.}$$

Ma se fosse colpito dopo 34 tiri: $I = 1 + \log_2(32/1) = 6$ bit, ritroveremmo tutta l'informazione.

Caccia al sottomarino

- Qual è la probabilità di colpire al primo colpo il sommergibile?
- Quanta informazione si otterrebbe?
- Si immagini di colpirlo dopo 1 o 2 o 63 colpi mancati come cambierebbe l'informazione ad ogni passo?
- $p_c = 1/64$ e $p_{nc} = 63/64$. Colpire subito significherebbe poter ricostruire tutta la tabella.
- $I = \log_2(64) = 6$ bit
- Ad ogni colpo l'informazione cambia, poichè la probabilità di trovare il sommergibile cambia. Se per esempio si eseguono 33 colpi mancati si ottiene un'informazione pari a:

$$I = \log_2(64/63) + \log_2(63/62) + \dots + \log_2(33/32) = 1 \text{ bit.}$$

Ma se fosse colpito dopo 34 tiri: $I = 1 + \log_2(32/1) = 6$ bit, ritroveremmo tutta l'informazione.

Caccia al sottomarino

- Qual è la probabilità di colpire al primo colpo il sommergibile?
- Quanta informazione si otterrebbe?
- Si immagini di colpirlo dopo 1 o 2 o 63 colpi mancati come cambierebbe l'informazione ad ogni passo?
- $p_c = 1/64$ e $p_{nc} = 63/64$. Colpire subito significherebbe poter ricostruire tutta la tabella.
- $I = \log_2(64) = 6$ bit
- Ad ogni colpo l'informazione cambia, poichè la probabilità di trovare il sommergibile cambia. Se per esempio si eseguono 33 colpi mancati si ottiene un'informazione pari a:

$$I = \log_2(64/63) + \log_2(63/62) + \dots + \log_2(33/32) = 1 \text{ bit.}$$

Ma se fosse colpito dopo 34 tiri: $I = 1 + \log_2(32/1) = 6$ bit, ritroveremmo tutta l'informazione.

Codifica di un messaggio

- Quante sono le lettere ed i simboli che ci permettono di scrivere in italiano?
- Se tu volessi usare $\{0, 1\}$ per codificare tali simboli, quanto dovranno essere lunghe le stringhe di 0 e 1?
- In un testo scritto in italiano, è vero che tutte le lettere capitano con la stessa frequenza?

Codifica di un messaggio

- Quante sono le lettere ed i simboli che ci permettono di scrivere in italiano?
- Se tu volessi usare $\{0, 1\}$ per codificare tali simboli, quanto dovranno essere lunghe le stringhe di 0 e 1?
- In un testo scritto in italiano, è vero che tutte le lettere capitano con la stessa frequenza?

Codifica di un messaggio

- Quante sono le lettere ed i simboli che ci permettono di scrivere in italiano?
- Se tu volessi usare $\{0, 1\}$ per codificare tali simboli, quanto dovranno essere lunghe le stringhe di 0 e 1?
- In un testo scritto in italiano, è vero che tutte le lettere capitano con la stessa frequenza?

Codifica di un messaggio

- Quante sono le lettere ed i simboli che ci permettono di scrivere in italiano?
- Se tu volessi usare $\{0, 1\}$ per codificare tali simboli, quanto dovranno essere lunghe le stringhe di 0 e 1?
- In un testo scritto in italiano, è vero che tutte le lettere capitano con la stessa frequenza?

La compressione

- I metodi di compressione dell'informazione: rappresentano i dati riducendo le ridondanze e rappresentare la stessa informazione con un numero inferiore di bit.
- Indichiamo con \mathcal{A} un alfabeto. È possibile rappresentare un $x \in \mathcal{A}$ con un numero di bit inferiori a quelli necessari in \mathcal{A} ?
- Far lavorare gli studenti su esempi semplici con alfabeti A con poche lettere.
- Risposta: No. Se questo fosse possibile vorrebbe dire che esisterebbe lunghezza l tale che

$$l < \log_2(|\mathcal{A}|)$$

e quindi $2^l < |\mathcal{A}|$. Questo comporterebbe che la rappresentazione non sarebbe iniettiva.

- Come funzionano allora gli algoritmi con i quali usualmente sono *zippati* i file ?

La compressione

- I metodi di compressione dell'informazione: rappresentano i dati riducendo le ridondanze e rappresentare la stessa informazione con un numero inferiore di bit.
- Indichiamo con \mathcal{A} un alfabeto. È possibile rappresentare un $x \in \mathcal{A}$ con un numero di bit inferiori a quelli necessari in \mathcal{A} ?
- Far lavorare gli studenti su esempi semplici con alfabeti A con poche lettere.
- Risposta: No. Se questo fosse possibile vorrebbe dire che esisterebbe lunghezza l tale che

$$l < \log_2(|\mathcal{A}|)$$

e quindi $2^l < |\mathcal{A}|$. Questo comporterebbe che la rappresentazione non sarebbe iniettiva.

- Come funzionano allora gli algoritmi con i quali usualmente sono *zippati* i file ?

La compressione

- I metodi di compressione dell'informazione: rappresentano i dati riducendo le ridondanze e rappresentare la stessa informazione con un numero inferiore di bit.
- Indichiamo con \mathcal{A} un alfabeto. È possibile rappresentare un $x \in \mathcal{A}$ con un numero di bit inferiori a quelli necessari in \mathcal{A} ?
- Far lavorare gli studenti su esempi semplici con alfabeti A con poche lettere.
- Risposta: No. Se questo fosse possibile vorrebbe dire che esisterebbe lunghezza l tale che

$$l < \log_2(|\mathcal{A}|)$$

e quindi $2^l < |\mathcal{A}|$. Questo comporterebbe che la rappresentazione non sarebbe iniettiva.

- Come funzionano allora gli algoritmi con i quali usualmente sono *zippati* i file ?

La compressione

- I metodi di compressione dell'informazione: rappresentano i dati riducendo le ridondanze e rappresentare la stessa informazione con un numero inferiore di bit.
- Indichiamo con \mathcal{A} un alfabeto. È possibile rappresentare un $x \in \mathcal{A}$ con un numero di bit inferiori a quelli necessari in \mathcal{A} ?
- Far lavorare gli studenti su esempi semplici con alfabeti A con poche lettere.
- Risposta: No. Se questo fosse possibile vorrebbe dire che esisterebbe lunghezza l tale che

$$l < \log_2(|\mathcal{A}|)$$

e quindi $2^l < |\mathcal{A}|$. Questo comporterebbe che la rappresentazione non sarebbe iniettiva.

- Come funzionano allora gli algoritmi con i quali usualmente sono *zippati* i file ?

La compressione

- I metodi di compressione dell'informazione: rappresentano i dati riducendo le ridondanze e rappresentare la stessa informazione con un numero inferiore di bit.
- Indichiamo con \mathcal{A} un alfabeto. È possibile rappresentare un $x \in \mathcal{A}$ con un numero di bit inferiori a quelli necessari in \mathcal{A} ?
- Far lavorare gli studenti su esempi semplici con alfabeti A con poche lettere.
- Risposta: No. Se questo fosse possibile vorrebbe dire che esisterebbe lunghezza l tale che

$$l < \log_2(|\mathcal{A}|)$$

e quindi $2^l < |\mathcal{A}|$. Questo comporterebbe che la rappresentazione non sarebbe iniettiva.

- Come funzionano allora gli algoritmi con i quali usualmente sono *zippati* i file ?

La compressione

- I metodi di compressione dell'informazione: rappresentano i dati riducendo le ridondanze e rappresentare la stessa informazione con un numero inferiore di bit.
- Indichiamo con \mathcal{A} un alfabeto. È possibile rappresentare un $x \in \mathcal{A}$ con un numero di bit inferiori a quelli necessari in \mathcal{A} ?
- Far lavorare gli studenti su esempi semplici con alfabeti A con poche lettere.
- Risposta: No. Se questo fosse possibile vorrebbe dire che esisterebbe lunghezza l tale che

$$l < \log_2(|\mathcal{A}|)$$

e quindi $2^l < |\mathcal{A}|$. Questo comporterebbe che la rappresentazione non sarebbe iniettiva.

- Come funzionano allora gli algoritmi con i quali usualmente sono *zippati* i file ?

Stringhe di lettere

- Consideriamo un alfabeto $\mathcal{A} = \{A, B, C\}$, ed una stringa s :

$$s = AAABCACBABACC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/13$, $f_B = 3/13$ e $f_C = 4/13$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2|\mathcal{A}| \simeq 2$, quindi 2 *bit*. Si può associare 0 a A , 01 a B e 11 a C .
- E quanti *bit* per la stringa s ? Ossia quanta informazione in *bit* è contenuta in s ?
- Potremmo calcolare la media del numero di *bit* utilizzati:

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

Stringhe di lettere

- Consideriamo un alfabeto $\mathcal{A} = \{A, B, C\}$, ed una stringa s :

$$s = AAABCACBABACC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/13$, $f_B = 3/13$ e $f_C = 4/13$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2|\mathcal{A}| \simeq 2$, quindi 2 *bit*. Si può associare 0 a A , 01 a B e 11 a C .
- E quanti *bit* per la stringa s ? Ossia quanta informazione in *bit* è contenuta in s ?
- Potremmo calcolare la media del numero di *bit* utilizzati:

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

Stringhe di lettere

- Consideriamo un alfabeto $\mathcal{A} = \{A, B, C\}$, ed una stringa s :

$$s = AAABCACBABACC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/13$, $f_B = 3/13$ e $f_C = 4/13$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2|\mathcal{A}| \simeq 2$, quindi 2 *bit*. Si può associare 0 a A, 01 a B e 11 a C
- E quanti *bit* per la stringa s ? Ossia quanta informazione in *bit* è contenuta in s ?
- Potremmo calcolare la media del numero di *bit* utilizzati:

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

Stringhe di lettere

- Consideriamo un alfabeto $\mathcal{A} = \{A, B, C\}$, ed una stringa s :

$$s = AAABCACBABACC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/13$, $f_B = 3/13$ e $f_C = 4/13$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2|\mathcal{A}| \simeq 2$, quindi 2 *bit*. Si può associare 0 a A, 01 a B e 11 a C
- E quanti *bit* per la stringa s ? Ossia quanta informazione in *bit* è contenuta in s ?
- Potremmo calcolare la media del numero di *bit* utilizzati:

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

Stringhe di lettere

- Consideriamo un alfabeto $\mathcal{A} = \{A, B, C\}$, ed una stringa s :

$$s = AAABCACBABACC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/13$, $f_B = 3/13$ e $f_C = 4/13$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2|\mathcal{A}| \simeq 2$, quindi 2 *bit*. Si può associare 0 a A, 01 a B e 11 a C
- E quanti *bit* per la stringa s ? Ossia quanta informazione in *bit* è contenuta in s ?
- Potremmo calcolare la media del numero di *bit* utilizzati:

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

Stringhe di lettere

- Consideriamo un alfabeto $\mathcal{A} = \{A, B, C\}$, ed una stringa s :

$$s = AAABCACBABACC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/13$, $f_B = 3/13$ e $f_C = 4/13$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2|\mathcal{A}| \simeq 2$, quindi 2 *bit*. Si può associare 0 a A, 01 a B e 11 a C
- E quanti *bit* per la stringa s ? Ossia quanta informazione in *bit* è contenuta in s ?
- Potremmo calcolare la media del numero di *bit* utilizzati:

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

Stringhe di lettere

- Consideriamo un alfabeto $\mathcal{A} = \{A, B, C\}$, ed una stringa s :

$$s = AAABCACBABACC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/13$, $f_B = 3/13$ e $f_C = 4/13$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2|\mathcal{A}| \simeq 2$, quindi 2 *bit*. Si può associare 0 a A, 01 a B e 11 a C
- E quanti *bit* per la stringa s ? Ossia quanta informazione in *bit* è contenuta in s ?
- Potremmo calcolare la media del numero di *bit* utilizzati:

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

Stringhe di lettere

- Si può generalizzare la formula?

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

- Consideriamo un altro esempio sempre con l'alfabeto $\mathcal{A} = \{A, B, C\}$, ma con una stringa s :

$$s = AAABCACBABAC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/12$, $f_B = 3/12$ e $f_C = 3/12$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2 |\mathcal{A}| \simeq 2$, quindi 2 *bit*. Si può associare 0 a A , 01 a B e 11 a C .
- Ossia quanta informazione in *bit* è contenuta in s ?
- Il numero medio del numero di *bit* utilizzati:

$$1 \frac{6}{12} + 2 \frac{3}{12} + 2 \frac{3}{12} = \frac{18}{12} = 1,50 \text{ bit}$$

Stringhe di lettere

- Si può generalizzare la formula?

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

- Consideriamo un altro esempio sempre con l'alfabeto $\mathcal{A} = \{A, B, C\}$, ma con una stringa s :

$$s = AAABCACBABAC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/12$, $f_B = 3/12$ e $f_C = 3/12$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2 |\mathcal{A}| \simeq 2$, quindi 2 *bit*. Si può associare 0 a A , 01 a B e 11 a C
- Ossia quanta informazione in *bit* è contenuta in s ?
- Il numero medio del numero di *bit* utilizzati:

$$1 \frac{6}{12} + 2 \frac{3}{12} + 2 \frac{3}{12} = \frac{18}{12} = 1,50 \text{ bit}$$

Stringhe di lettere

- Si può generalizzare la formula?

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

- Consideriamo un altro esempio sempre con l'alfabeto $\mathcal{A} = \{A, B, C\}$, ma con una stringa s :

$$s = AAABCACBABAC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/12$, $f_B = 3/12$ e $f_C = 3/12$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2 |\mathcal{A}| \simeq 2$, quindi 2 *bit*. Si può associare 0 a A , 01 a B e 11 a C
- Ossia quanta informazione in *bit* è contenuta in s ?
- Il numero medio del numero di *bit* utilizzati:

$$1 \frac{6}{12} + 2 \frac{3}{12} + 2 \frac{3}{12} = \frac{18}{12} = 1,50 \text{ bit}$$

Stringhe di lettere

- Si può generalizzare la formula?

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

- Consideriamo un altro esempio sempre con l'alfabeto $\mathcal{A} = \{A, B, C\}$, ma con una stringa s :

$$s = AAABCACBABAC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/12$, $f_B = 3/12$ e $f_C = 3/12$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2 |\mathcal{A}| \simeq 2$, quindi *2 bit*. Si può associare 0 a *A*, 01 a *B* e 11 a *C*
- Ossia quanta informazione in *bit* è contenuta in s ?
- Il numero medio del numero di *bit* utilizzati:

$$1 \frac{6}{12} + 2 \frac{3}{12} + 2 \frac{3}{12} = \frac{18}{12} = 1,50 \text{ bit}$$

Stringhe di lettere

- Si può generalizzare la formula?

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

- Consideriamo un altro esempio sempre con l'alfabeto $\mathcal{A} = \{A, B, C\}$, ma con una stringa s :

$$s = AAABCACBABAC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/12$, $f_B = 3/12$ e $f_C = 3/12$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2 |\mathcal{A}| \simeq 2$, quindi *2 bit*. Si può associare 0 a *A*, 01 a *B* e 11 a *C*
- Ossia quanta informazione in *bit* è contenuta in s ?
- Il numero medio del numero di *bit* utilizzati:

$$1 \frac{6}{12} + 2 \frac{3}{12} + 2 \frac{3}{12} = \frac{18}{12} = 1,50 \text{ bit}$$

Stringhe di lettere

- Si può generalizzare la formula?

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

- Consideriamo un altro esempio sempre con l'alfabeto $\mathcal{A} = \{A, B, C\}$, ma con una stringa s :

$$s = AAABCACBABAC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/12$, $f_B = 3/12$ e $f_C = 3/12$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2 |\mathcal{A}| \simeq 2$, quindi *2 bit*. Si può associare 0 a *A*, 01 a *B* e 11 a *C*
- Ossia quanta informazione in *bit* è contenuta in s ?
- Il numero medio del numero di *bit* utilizzati:

$$1 \frac{6}{12} + 2 \frac{3}{12} + 2 \frac{3}{12} = \frac{18}{12} = 1,50 \text{ bit}$$

Stringhe di lettere

- Si può generalizzare la formula?

$$1 \frac{6}{13} + 2 \frac{3}{13} + 2 \frac{4}{13} = \frac{20}{13} = 1,53 \text{ bit}$$

- Consideriamo un altro esempio sempre con l'alfabeto $\mathcal{A} = \{A, B, C\}$, ma con una stringa s :

$$s = AAABCACBABAC.$$

- Con quale frequenza compaiono le varie lettere?
- $f_A = 6/12$, $f_B = 3/12$ e $f_C = 3/12$.
- Quanti *bit* utilizzeresti per codificare le singole lettere di \mathcal{A} ?
- $\log_2 |\mathcal{A}| \simeq 2$, quindi *2 bit*. Si può associare 0 a *A*, 01 a *B* e 11 a *C*
- Ossia quanta informazione in *bit* è contenuta in s ?
- Il numero medio del numero di *bit* utilizzati:

$$1 \frac{6}{12} + 2 \frac{3}{12} + 2 \frac{3}{12} = \frac{18}{12} = 1,50 \text{ bit}$$

A caccia di una formula...

- Ora la formula

$$1 \frac{6}{12} + 2 \frac{3}{12} + 2 \frac{3}{12} = \frac{18}{12} = 1,50 \text{ bit}$$

può suggerire una sua generalizzazione...

-

$$1 \frac{1}{2} + 2 \frac{1}{4} + 2 \frac{1}{4} = \frac{6}{4} = 1,50 \text{ bit}$$

- Ricordiamo che l'informazione può essere valutata con

$$\log_2(1/p) = -\log_2(p)$$

E allora...

-

$$-\frac{1}{2} \log_2(2) - \frac{1}{4} \log_2(4) - \frac{1}{4} \log_2(4) = 1,5 \text{ bit}$$

A caccia di una formula...

- Ora la formula

$$1 \frac{6}{12} + 2 \frac{3}{12} + 2 \frac{3}{12} = \frac{18}{12} = 1,50 \text{ bit}$$

può suggerire una sua generalizzazione...

-

$$1 \frac{1}{2} + 2 \frac{1}{4} + 2 \frac{1}{4} = \frac{6}{4} = 1,50 \text{ bit}$$

- Ricordiamo che l'informazione può essere valutata con

$$\log_2(1/p) = -\log_2(p)$$

E allora...

-

$$-\frac{1}{2} \log_2(2) - \frac{1}{4} \log_2(4) - \frac{1}{4} \log_2(4) = 1,5 \text{ bit}$$

A caccia di una formula...

- Ora la formula

$$1 \frac{6}{12} + 2 \frac{3}{12} + 2 \frac{3}{12} = \frac{18}{12} = 1,50 \text{ bit}$$

può suggerire una sua generalizzazione...

-

$$1 \frac{1}{2} + 2 \frac{1}{4} + 2 \frac{1}{4} = \frac{6}{4} = 1,50 \text{ bit}$$

- Ricordiamo che l'informazione può essere valutata con

$$\log_2(1/p) = -\log_2(p)$$

E allora...

-

$$-\frac{1}{2} \log_2(2) - \frac{1}{4} \log_2(4) - \frac{1}{4} \log_2(4) = 1,5 \text{ bit}$$

A caccia di una formula...

- Ora la formula

$$1 \frac{6}{12} + 2 \frac{3}{12} + 2 \frac{3}{12} = \frac{18}{12} = 1,50 \text{ bit}$$

può suggerire una sua generalizzazione...

-

$$1 \frac{1}{2} + 2 \frac{1}{4} + 2 \frac{1}{4} = \frac{6}{4} = 1,50 \text{ bit}$$

- Ricordiamo che l'informazione può essere valutata con

$$\log_2(1/p) = -\log_2(p)$$

E allora...

-

$$-\frac{1}{2} \log_2(2) - \frac{1}{4} \log_2(4) - \frac{1}{4} \log_2(4) = 1,5 \text{ bit}$$

A caccia di una formula...

- La formula

$$-\frac{1}{2} \log_2(2) - \frac{1}{4} \log_2(4) - \frac{1}{4} \log_2(4) = 1,5 \text{ bit}$$

si può interpretare come:

-

$$-f_A \log_2(1/f_A) - f_B \log_2(1/f_B) - f_C \log_2(1/f_C)$$

A caccia di una formula...

- La formula

$$-\frac{1}{2} \log_2(2) - \frac{1}{4} \log_2(4) - \frac{1}{4} \log_2(4) = 1,5 \text{ bit}$$

si può interpretare come:

-

$$-f_A \log_2(1/f_A) - f_B \log_2(1/f_B) - f_C \log_2(1/f_C)$$

A caccia di una formula...

- La formula

$$-\frac{1}{2} \log_2(2) - \frac{1}{4} \log_2(4) - \frac{1}{4} \log_2(4) = 1,5 \text{ bit}$$

si può interpretare come:

-

$$-f_A \log_2(1/f_A) - f_B \log_2(1/f_B) - f_C \log_2(1/f_C)$$

Entropia di Shannon e stringhe tipiche

Consideriamo alfabeto $\mathcal{A} = \{x_1, \dots, x_a\}$ ed una lingua in cui ciascun x_i compare con una probabilità caratteristica p_i .

Prendiamo s è una stringa di lunghezza N di elementi di \mathcal{A} :

- Quante volte potrà capitare x_i in s ?
- Ci aspettiamo un numero pari a Np_i .
- Quale sarà quindi la probabilità di avere la stringa s ?
- La scelta delle lettere possiamo assumerla una collezione di eventi indipendenti: $p(s) = p_1^{Np_1} p_2^{Np_2} \dots p_a^{Np_a}$.
- Quanta informazione conterrà s ?
- $I(s) = 1/\log_2(P(s)) = -\log_2(P(s)) = -N \sum_{i=1}^a p_i \log_2(p_i)$.
L'espressione $H_S(p_1, \dots, p_a) = -\sum_{i=1}^a p_i \log_2(p_i)$ è detta entropia di Shannon.

Entropia di Shannon e stringhe tipiche

Consideriamo alfabeto $\mathcal{A} = \{x_1, \dots, x_a\}$ ed una lingua in cui ciascun x_j compare con una probabilità caratteristica p_j .

Prendiamo s è una stringa di lunghezza N di elementi di \mathcal{A} :

- Quante volte potrà capitare x_j in s ?
- Ci aspettiamo un numero pari a Np_j .
- Quale sarà quindi la probabilità di avere la stringa s ?
- La scelta delle lettere possiamo assumerla una collezione di eventi indipendenti: $p(s) = p_1^{Np_1} p_2^{Np_2} \dots p_a^{Np_a}$.
- Quanta informazione conterrà s ?
- $I(s) = 1/\log_2(P(s)) = -\log_2(P(s)) = -N \sum_{i=1}^a p_i \log_2(p_i)$.
L'espressione $H_S(p_1, \dots, p_a) = -\sum_{i=1}^a p_i \log_2(p_i)$ è detta entropia di Shannon.

Entropia di Shannon e stringhe tipiche

Consideriamo alfabeto $\mathcal{A} = \{x_1, \dots, x_a\}$ ed una lingua in cui ciascun x_j compare con una probabilità caratteristica p_j .

Prendiamo s è una stringa di lunghezza N di elementi di \mathcal{A} :

- Quante volte potrà capitare x_j in s ?
- **Ci aspettiamo un numero pari a Np_j .**
- Quale sarà quindi la probabilità di avere la stringa s ?
- La scelta delle lettere possiamo assumerla una collezione di eventi indipendenti: $p(s) = p_1^{Np_1} p_2^{Np_2} \dots p_a^{Np_a}$.
- Quanta informazione conterrà s ?
- $I(s) = 1/\log_2(P(s)) = -\log_2(P(s)) = -N \sum_{i=1}^a p_i \log_2(p_i)$.
L'espressione $H_S(p_1, \dots, p_a) = -\sum_{i=1}^a p_i \log_2(p_i)$ è detta entropia di Shannon.

Entropia di Shannon e stringhe tipiche

Consideriamo alfabeto $\mathcal{A} = \{x_1, \dots, x_a\}$ ed una lingua in cui ciascun x_j compare con una probabilità caratteristica p_j .

Prendiamo s è una stringa di lunghezza N di elementi di \mathcal{A} :

- Quante volte potrà capitare x_j in s ?
- **Ci aspettiamo un numero pari a Np_j .**
- Quale sarà quindi la probabilità di avere la stringa s ?
- La scelta delle lettere possiamo assumerla una collezione di eventi indipendenti: $p(s) = p_1^{Np_1} p_2^{Np_2} \dots p_a^{Np_a}$.
- Quanta informazione conterrà s ?
- $I(s) = 1/\log_2(P(s)) = -\log_2(P(s)) = -N \sum_{i=1}^a p_i \log_2(p_i)$.
L'espressione $H_S(p_1, \dots, p_a) = -\sum_{i=1}^a p_i \log_2(p_i)$ è detta entropia di Shannon.

Entropia di Shannon e stringhe tipiche

Consideriamo alfabeto $\mathcal{A} = \{x_1, \dots, x_a\}$ ed una lingua in cui ciascun x_j compare con una probabilità caratteristica p_j .

Prendiamo s è una stringa di lunghezza N di elementi di \mathcal{A} :

- Quante volte potrà capitare x_j in s ?
- Ci aspettiamo un numero pari a Np_j .
- Quale sarà quindi la probabilità di avere la stringa s ?
- La scelta delle lettere possiamo assumerla una collezione di eventi indipendenti: $p(s) = p_1^{Np_1} p_2^{Np_2} \dots p_a^{Np_a}$.
- Quanta informazione conterrà s ?
- $I(s) = 1/\log_2(P(s)) = -\log_2(P(s)) = -N \sum_{i=1}^a p_i \log_2(p_i)$.
L'espressione $H_S(p_1, \dots, p_a) = -\sum_{i=1}^a p_i \log_2(p_i)$ è detta entropia di Shannon.

Entropia di Shannon e stringhe tipiche

Consideriamo alfabeto $\mathcal{A} = \{x_1, \dots, x_a\}$ ed una lingua in cui ciascun x_j compare con una probabilità caratteristica p_j .

Prendiamo s è una stringa di lunghezza N di elementi di \mathcal{A} :

- Quante volte potrà capitare x_j in s ?
- Ci aspettiamo un numero pari a Np_j .
- Quale sarà quindi la probabilità di avere la stringa s ?
- La scelta delle lettere possiamo assumerla una collezione di eventi indipendenti: $p(s) = p_1^{Np_1} p_2^{Np_2} \dots p_a^{Np_a}$.
- Quanta informazione conterrà s ?
- $I(s) = 1/\log_2(P(s)) = -\log_2(P(s)) = -N \sum_{i=1}^a p_i \log_2(p_i)$.
L'espressione $H_S(p_1, \dots, p_a) = -\sum_{i=1}^a p_i \log_2(p_i)$ è detta entropia di Shannon.

Entropia di Shannon e stringhe tipiche

Consideriamo alfabeto $\mathcal{A} = \{x_1, \dots, x_a\}$ ed una lingua in cui ciascun x_j compare con una probabilità caratteristica p_j .

Prendiamo s è una stringa di lunghezza N di elementi di \mathcal{A} :

- Quante volte potrà capitare x_j in s ?
- Ci aspettiamo un numero pari a Np_j .
- Quale sarà quindi la probabilità di avere la stringa s ?
- La scelta delle lettere possiamo assumerla una collezione di eventi indipendenti: $p(s) = p_1^{Np_1} p_2^{Np_2} \dots p_a^{Np_a}$.
- Quanta informazione conterrà s ?
- $I(s) = 1/\log_2(P(s)) = -\log_2(P(s)) = -N \sum_{i=1}^a p_i \log_2(p_i)$.
L'espressione $H_S(p_1, \dots, p_a) = -\sum_{i=1}^a p_i \log_2(p_i)$ è detta entropia di Shannon.

Proprietà dell'entropia di Shannon

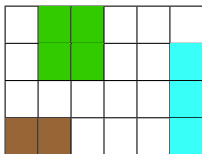
L'entropia di Shannon permette di valutare il minimo del numero medio di *bit* che occorrono per codificare una stringa.

Ordine e disordine: l'entropia di Gibbs

- Immagina la tua scrivania o la tua stanza: certamente ci sono un certo numero N d'oggetti che arredano la stanza e che sono tuo uso comune. Questi oggetti possono evidentemente essere disposti in altrettante posizioni.
- Se tu assumessi di poterli disporre a tuo assoluto piacimento...ovvero *senza un criterio*...quale sarebbe la probabilità con la quale una penna sia fra le camicie?
- Tua madre entra in camera...e...visibilmente alterata, ti intima di mettere ordine. Qual è la differenza fra un camera *ordinata* ed una *disordinata*?
- Per esempio se la camera è *in ordine*, è più o meno probabile trovare un libro nella libreria che sul tappeto?
- Qual è la differenza fra uno stato *ordinato* della tua camera ed uno *disordinato*?

Scheda: ordine e disordine nella *mia* stanza

j



- La tua stanza è suddivisa può essere schematizzata con un quadrato di 24 quadretti.
- I 4 quadretti verdi indicano la scrivania, i 2 quadretti marroni sono la libreria e i 3 azzurri, l'armadio.
- La tua stanza come sistema fisico può essere descritto da due stati: $S_1 = \text{ordinata}$ e $S_2 = \text{disordinata}$.
- Immagina di avere una penna P , un libro L ed una camicia C .

Scheda: ordine e disordine nella *mia* stanza

- Usando la scheda disponi P , L e C in modoragionevolmente ordinato $S_1... :-)$ Quante possibili posizioni puoi scegliere per *ciascun* oggetto mantenendo la stanza ordinata?
- Usando la seconda scheda disponi P , L e C in mododisordinato $S_2... :-/$ Quante possibili posizioni puoi scegliere per *ciascun* oggetto mantenendo la stanza disordinata?
Sei autorizzato ad essere veramente disordinato quindi potresti, per esempio, mettere la penna e la camicia nello stesso riquadro!!!
- Gli stati S_1 e S_2 della stanza sono due esempi di **stati macroscopici**.
- Gli stati invece di L , P e C sono invece detti **stati microscopici**.

Le probabilità

Le probabilità che troverai saranno:

- $\mathbb{P}(\text{Penna sulla scrivania}) = 4/24 = 1/6$
- $\mathbb{P}(\text{Libro sulla libreria}) = 2/24 = 1/12$
- $\mathbb{P}(\text{Camicia nell'armadio}) = 3/24 = 1/8$
- $\mathbb{P}(\text{Un oggetto è in un qualunque punto fissato della stanza}) = 1/24$

Le probabilità

Le probabilità che troverai saranno:

- $\mathbb{P}(\text{Penna sulla scrivania}) = 4/24 = 1/6$
- $\mathbb{P}(\text{Libro sulla libreria}) = 2/24 = 1/12$
- $\mathbb{P}(\text{Camicia nell'armadio}) = 3/24 = 1/8$
- $\mathbb{P}(\text{Un oggetto è in un qualunque punto fissato della stanza}) = 1/24$

Le probabilità

Le probabilità che troverai saranno:

- $\mathbb{P}(\text{Penna sulla scrivania}) = 4/24 = 1/6$
- $\mathbb{P}(\text{Libro sulla libreria}) = 2/24 = 1/12$
- $\mathbb{P}(\text{Camicia nell'armadio}) = 3/24 = 1/8$
- $\mathbb{P}(\text{Un oggetto è in un qualunque punto fissato della stanza}) = 1/24$

Le probabilità

Le probabilità che troverai saranno:

- $\mathbb{P}(\text{Penna sulla scrivania}) = 4/24 = 1/6$
- $\mathbb{P}(\text{Libro sulla libreria}) = 2/24 = 1/12$
- $\mathbb{P}(\text{Camicia nell'armadio}) = 3/24 = 1/8$
- $\mathbb{P}(\text{Un oggetto è in un qualunque punto fissato della stanza}) = 1/24$

Le probabilità

Le probabilità che troverai saranno:

- $\mathbb{P}(\text{Penna sulla scrivania}) = 4/24 = 1/6$
- $\mathbb{P}(\text{Libro sulla libreria}) = 2/24 = 1/12$
- $\mathbb{P}(\text{Camicia nell'armadio}) = 3/24 = 1/8$
- $\mathbb{P}(\text{Un oggetto è in un qualunque punto fissato della stanza}) = 1/24$

L'entropia di Gibbs

- Supponiamo che un sistema fisico sia formato da N elementi. Lo stato S del sistema sarà quindi definito sapendo lo stato di ciascuno degli N oggetti: lo stato di *ordine* o *disordine* della tua camera dipende dallo stato degli oggetti presenti.
- Ogni oggetto componente il sistema può essere in un suo stato detto *microstato* con una probabilità p_i .
- Si chiama entropia di Gibbs la seguente espressione:

$$H_G(S) = H_G(p_1, \dots, p_N) = -k_B \sum_{i=1}^N p_i \ln(p_i).$$

Calcolo dell'entropia di Gibbs



$$H_G(p_1, \dots, p_N) = - \sum_{i=1}^N p_i \ln(p_i)$$



$$H_G(1/N, \dots, 1/N) = - \sum_{i=1}^N \frac{1}{N} \ln\left(\frac{1}{N}\right)$$

- Calcoliamo la variazione dell'entropia:

$$\Delta H_G = H_G(1/N, \dots, 1/N) - H_G(p_1, \dots, p_N)$$

$$\Delta H_G = \sum_{i=1}^N \left[p_i \ln(p_i) - \frac{1}{N} \ln(p_i) + \frac{1}{N} \ln(p_i) - \frac{1}{N} \ln\left(\frac{1}{N}\right) \right]$$

$$\Delta H_G = \sum_{i=1}^N \left[\left(p_i - \frac{1}{N} \right) \ln(p_i) + \frac{1}{N} \ln(N p_i) \right]$$

Calcolo dell'entropia di Gibbs



$$H_G(p_1, \dots, p_N) = - \sum_{i=1}^N p_i \ln(p_i)$$



$$H_G(1/N, \dots, 1/N) = - \sum_{i=1}^N \frac{1}{N} \ln\left(\frac{1}{N}\right)$$

- Calcoliamo la variazione dell'entropia:

$$\Delta H_G = H_G(1/N, \dots, 1/N) - H_G(p_1, \dots, p_N)$$

$$\Delta H_G = \sum_{i=1}^N \left[p_i \ln(p_i) - \frac{1}{N} \ln(p_i) + \frac{1}{N} \ln(p_i) - \frac{1}{N} \ln\left(\frac{1}{N}\right) \right]$$

$$\Delta H_G = \sum_{i=1}^N \left[\left(p_i - \frac{1}{N} \right) \ln(p_i) + \frac{1}{N} \ln(N p_i) \right]$$

Calcolo dell'entropia di Gibbs



$$H_G(p_1, \dots, p_N) = - \sum_{i=1}^N p_i \ln(p_i)$$



$$H_G(1/N, \dots, 1/N) = - \sum_{i=1}^N \frac{1}{N} \ln\left(\frac{1}{N}\right)$$

- Calcoliamo la variazione dell'entropia:

$$\Delta H_G = H_G(1/N, \dots, 1/N) - H_G(p_1, \dots, p_N)$$

$$\Delta H_G = \sum_{i=1}^N \left[p_i \ln(p_i) - \frac{1}{N} \ln(p_i) + \frac{1}{N} \ln(p_i) - \frac{1}{N} \ln\left(\frac{1}{N}\right) \right]$$

$$\Delta H_G = \sum_{i=1}^N \left[\left(p_i - \frac{1}{N} \right) \ln(p_i) + \frac{1}{N} \ln(N p_i) \right]$$

Calcolo dell'entropia di Gibbs



$$H_G(p_1, \dots, p_N) = - \sum_{i=1}^N p_i \ln(p_i)$$



$$H_G(1/N, \dots, 1/N) = - \sum_{i=1}^N \frac{1}{N} \ln\left(\frac{1}{N}\right)$$

- Calcoliamo la variazione dell'entropia:

$$\Delta H_G = H_G(1/N, \dots, 1/N) - H_G(p_1, \dots, p_N)$$

$$\Delta H_G = \sum_{i=1}^N \left[p_i \ln(p_i) - \frac{1}{N} \ln(p_i) + \frac{1}{N} \ln(p_i) - \frac{1}{N} \ln\left(\frac{1}{N}\right) \right]$$

$$\Delta H_G = \sum_{i=1}^N \left[\left(p_i - \frac{1}{N} \right) \ln(p_i) + \frac{1}{N} \ln(N p_i) \right]$$

Calcolo dell'entropia di Gibbs

- L'entropia dello stato in cui si ha la *stanza ordinata*:

$$H(S_1) = - \left[\frac{1}{6} \ln \left(\frac{1}{6} \right) + \frac{1}{12} \ln \left(\frac{1}{12} \right) + \frac{1}{8} \ln \left(\frac{1}{8} \right) \right]$$

ovvero

$$H(S_1) = \frac{3}{12} \ln(3) + \frac{17}{24} \ln(2).$$

- L'entropia dello stato in cui si ha la *stanza disordinata*:

$$H(S_2) = - \sum_{i=1}^{24} \frac{1}{24} \ln \left(\frac{1}{24} \right) = \ln(3) + 3 \ln(2)$$

- Da cui si verifica facilmente:

$$H(S_1) < H(S_2)$$

Calcolo dell'entropia di Gibbs

- L'entropia dello stato in cui si ha la *stanza ordinata*:

$$H(S_1) = - \left[\frac{1}{6} \ln \left(\frac{1}{6} \right) + \frac{1}{12} \ln \left(\frac{1}{12} \right) + \frac{1}{8} \ln \left(\frac{1}{8} \right) \right]$$

ovvero

$$H(S_1) = \frac{3}{12} \ln(3) + \frac{17}{24} \ln(2).$$

- L'entropia dello stato in cui si ha la *stanza disordinata*:

$$H(S_2) = - \sum_{i=1}^{24} \frac{1}{24} \ln \left(\frac{1}{24} \right) = \ln(3) + 3 \ln(2)$$

- Da cui si verifica facilmente:

$$H(S_1) < H(S_2)$$

Calcolo dell'entropia di Gibbs

- L'entropia dello stato in cui si ha la *stanza ordinata*:

$$H(S_1) = - \left[\frac{1}{6} \ln \left(\frac{1}{6} \right) + \frac{1}{12} \ln \left(\frac{1}{12} \right) + \frac{1}{8} \ln \left(\frac{1}{8} \right) \right]$$

ovvero

$$H(S_1) = \frac{3}{12} \ln(3) + \frac{17}{24} \ln(2).$$

- L'entropia dello stato in cui si ha la *stanza disordinata*:

$$H(S_2) = - \sum_{i=1}^{24} \frac{1}{24} \ln \left(\frac{1}{24} \right) = \ln(3) + 3 \ln(2)$$

- Da cui si verifica facilmente:

$$H(S_1) < H(S_2)$$

Calcolo dell'entropia di Gibbs

- L'entropia dello stato in cui si ha la *stanza ordinata*:

$$H(S_1) = - \left[\frac{1}{6} \ln \left(\frac{1}{6} \right) + \frac{1}{12} \ln \left(\frac{1}{12} \right) + \frac{1}{8} \ln \left(\frac{1}{8} \right) \right]$$

ovvero

$$H(S_1) = \frac{3}{12} \ln(3) + \frac{17}{24} \ln(2).$$

- L'entropia dello stato in cui si ha la *stanza disordinata*:

$$H(S_2) = - \sum_{i=1}^{24} \frac{1}{24} \ln \left(\frac{1}{24} \right) = \ln(3) + 3 \ln(2)$$

- Da cui si verifica facilmente:

$$H(S_1) < H(S_2)$$





L'entropia di Boltzman

- Consideriamo un gas in un contenitore. Da cosa è formato?
- **Atomi o molecole**
- Usualmente si misura un gas in litri o meglio in moli. Qual è l'ordine di grandezza delle particelle che dovremo considerare?
- **Circa 10^{23} particelle.**
- Come possiamo definire lo stato meccanico di ogni singola particella?
- **Dal punto di vista della meccanica lo stato è definito se si determinassero posizione \vec{x} e velocità, o meglio quantità di moto \vec{p} per ogni particella. (MICROSTATO: Definiamo lo stato di una singola particella la coppia (\vec{x}, \vec{p}) data ad un tempo t .)**

L'entropia di Boltzman

- Consideriamo un gas in un contenitore. Da cosa è formato?
- Atomi o molecole
- Usualmente si misura un gas in litri o meglio in moli. Qual è l'ordine di grandezza delle particelle che dovremo considerare?
- Circa 10^{23} particelle.
- Come possiamo definire lo stato meccanico di ogni singola particella?
- Dal punto di vista della meccanica lo stato è definito se si determinassero posizione \vec{x} e velocità, o meglio quantità di moto \vec{p} per ogni particella. (MICROSTATO: Definiamo lo stato di una singola particella la coppia (\vec{x}, \vec{p}) data ad un tempo t .)

L'entropia di Boltzman

- Consideriamo un gas in un contenitore. Da cosa è formato?
- **Atomi o molecole**
- Usualmente si misura un gas in litri o meglio in moli. Qual è l'ordine di grandezza delle particelle che dovremo considerare?
- Circa 10^{23} particelle.
- Come possiamo definire lo stato meccanico di ogni singola particella?
- Dal punto di vista della meccanica lo stato è definito se si determinassero posizione \vec{x} e velocità, o meglio quantità di moto \vec{p} per ogni particella. (MICROSTATO: Definiamo lo stato di una singola particella la coppia (\vec{x}, \vec{p}) data ad un tempo t .)

L'entropia di Boltzman

- Consideriamo un gas in un contenitore. Da cosa è formato?
- **Atomi o molecole**
- Usualmente si misura un gas in litri o meglio in moli. Qual è l'ordine di grandezza delle particelle che dovremo considerare?
- Circa 10^{23} particelle.
- Come possiamo definire lo stato meccanico di ogni singola particella?
- Dal punto di vista della meccanica lo stato è definito se si determinassero posizione \vec{x} e velocità, o meglio quantità di moto \vec{p} per ogni particella. (MICROSTATO: Definiamo lo stato di una singola particella la coppia (\vec{x}, \vec{p}) data ad un tempo t .)

L'entropia di Boltzman

- Consideriamo un gas in un contenitore. Da cosa è formato?
- **Atomi o molecole**
- Usualmente si misura un gas in litri o meglio in moli. Qual è l'ordine di grandezza delle particelle che dovremo considerare?
- **Circa 10^{23} particelle.**
- Come possiamo definire lo stato meccanico di ogni singola particella?
- Dal punto di vista della meccanica lo stato è definito se si determinassero posizione \vec{x} e velocità, o meglio quantità di moto \vec{p} per ogni particella. (MICROSTATO: Definiamo lo stato di una singola particella la coppia (\vec{x}, \vec{p}) data ad un tempo t .)

L'entropia di Boltzman

- Consideriamo un gas in un contenitore. Da cosa è formato?
- **Atomi o molecole**
- Usualmente si misura un gas in litri o meglio in moli. Qual è l'ordine di grandezza delle particelle che dovremo considerare?
- **Circa 10^{23} particelle.**
- Come possiamo definire lo stato meccanico di ogni singola particella?
- Dal punto di vista della meccanica lo stato è definito se si determinassero posizione \vec{x} e velocità, o meglio quantità di moto \vec{p} per ogni particella. (MICROSTATO: Definiamo lo stato di una singola particella la coppia (\vec{x}, \vec{p}) data ad un tempo t .)

L'entropia di Boltzman

- Consideriamo un gas in un contenitore. Da cosa è formato?
- **Atomi o molecole**
- Usualmente si misura un gas in litri o meglio in moli. Qual è l'ordine di grandezza delle particelle che dovremo considerare?
- **Circa 10^{23} particelle.**
- Come possiamo definire lo stato meccanico di ogni singola particella?
- **Dal punto di vista della meccanica lo stato è definito se si determinassero posizione \vec{x} e velocità, o meglio quantità di moto \vec{p} per ogni particella. (MICROSTATO: Definiamo lo stato di una singola particella la coppia (\vec{x}, \vec{p}) data ad un tempo t .)**

L'entropia di Boltzman

- Ora supponiamo di assegnare un'energia E al gas, quali conseguenze ci saranno per le singole particelle? Puoi immaginare esempi concreti che ti guidino nella riflessione?
- Ciascuna particella riceverà una parte dell'energia e si metterà in moto ed urterà contro le pareti del contenitore.
- Potremmo modificare lo stato di una mole di gas se modificassimo lo stato di singola particella?

L'entropia di Boltzman

- Ora supponiamo di assegnare un'energia E al gas, quali conseguenze ci saranno per le singole particelle? Puoi immaginare esempi concreti che ti guidino nella riflessione?
- Ciascuna particella riceverà una parte dell'energia e si metterà in moto ed urterà contro le pareti del contenitore.
- Potremmo modificare lo stato di una mole di gas se modificassimo lo stato di singola particella?

L'entropia di Boltzman

- Ora supponiamo di assegnare un'energia E al gas, quali conseguenze ci saranno per le singole particelle? Puoi immaginare esempi concreti che ti guidino nella riflessione?
- Ciascuna particella riceverà una parte dell'energia e si metterà in moto ed urterà contro le pareti del contenitore.
- Potremmo modificare lo stato di una mole di gas se modificassimo lo stato di singola particella?

L'entropia di Boltzman

- Ora supponiamo di assegnare un'energia E al gas, quali conseguenze ci saranno per le singole particelle? Puoi immaginare esempi concreti che ti guidino nella riflessione?
- Ciascuna particella riceverà una parte dell'energia e si metterà in moto ed urterà contro le pareti del contenitore.
- Potremmo modificare lo stato di una mole di gas se modificassimo lo stato di singola particella?

L'entropia di Boltzman

- Potremmo modificare lo stato di una mole di gas se modificassimo lo stato di singola particella?
- In effetti nessuno. L.Boltzman ai primi del 1900 mostrò che se consideriamo:
 - E energia complessiva delle particelle di un gas,
 - $\Omega(E)$ volume di tutti i possibili microstati permessi alle particelle del gas,
 - La funzione

$$H_B(E) = k_B \ln(\Omega(E))$$

L'entropia di Boltzman

- Potremmo modificare lo stato di una mole di gas se modificassimo lo stato di singola particella?
- In effetti nessuno. L.Boltzman ai primi del 1900 mostrò che se consideriamo:
 - E energia complessiva delle particelle di un gas,
 - $\Omega(E)$ volume di tutti i possibili microstati permessi alle particelle del gas,
 - La funzione

$$H_B(E) = k_B \ln(\Omega(E))$$

L'entropia di Boltzman

- A partire dalla formula:

$$S(E) = k_B \ln(\Omega(E))$$

Osserviamo:

- Se aumentiamo E come cambia il volume dei possibili stati delle particelle $\Omega(E)$?
- Cresce se E cresce, decresce se E diminuisce.
- Corrispondentemente come cambierà l'entropia $S(E)$?
- Rivedersi le proprietà dell'esponenziale...
- Boltzman dimostrò che $S(E)$ è equivalente alla entropia termodinamica ed in particolare che la temperatura del gas poteva essere calcolata:

$$\frac{1}{T} = \frac{\Delta S(E)}{\Delta E}$$

L'entropia di Boltzman

- A partire dalla formula:

$$S(E) = k_B \ln(\Omega(E))$$

Osserviamo:

- Se aumentiamo E come cambia il volume dei possibili stati delle particelle $\Omega(E)$?
- Cresce se E cresce, decresce se E diminuisce.
- Corrispondentemente come cambierà l'entropia $S(E)$?
- Rivedersi le proprietà dell'esponenziale...
- Boltzman dimostrò che $S(E)$ è equivalente alla entropia termodinamica ed in particolare che la temperatura del gas poteva essere calcolata:

$$\frac{1}{T} = \frac{\Delta S(E)}{\Delta E}$$

L'entropia di Boltzman

- A partire dalla formula:

$$S(E) = k_B \ln(\Omega(E))$$

Osserviamo:

- Se aumentiamo E come cambia il volume dei possibili stati delle particelle $\Omega(E)$?
- **Cresce se E cresce, decresce se E diminuisce.**
- Corrispondentemente come cambierà l'entropia $S(E)$?
- **Rivedersi le proprietà dell'esponenziale...**
- Boltzman dimostrò che $S(E)$ è equivalente alla entropia termodinamica ed in particolare che la temperatura del gas poteva essere calcolata:

$$\frac{1}{T} = \frac{\Delta S(E)}{\Delta E}$$

L'entropia di Boltzman

- A partire dalla formula:

$$S(E) = k_B \ln(\Omega(E))$$

Osserviamo:

- Se aumentiamo E come cambia il volume dei possibili stati delle particelle $\Omega(E)$?
- **Cresce se E cresce, decresce se E diminuisce.**
- Corrispondentemente come cambierà l'entropia $S(E)$?
- **Rivedersi le proprietà dell'esponenziale...**
- Boltzman dimostrò che $S(E)$ è equivalente alla entropia termodinamica ed in particolare che la temperatura del gas poteva essere calcolata:

$$\frac{1}{T} = \frac{\Delta S(E)}{\Delta E}$$

L'entropia di Boltzman

- A partire dalla formula:

$$S(E) = k_B \ln(\Omega(E))$$

Osserviamo:

- Se aumentiamo E come cambia il volume dei possibili stati delle particelle $\Omega(E)$?
- **Cresce se E cresce, decresce se E diminuisce.**
- Corrispondentemente come cambierà l'entropia $S(E)$?
- **Rivedersi le proprietà dell'esponenziale...**
- Boltzman dimostrò che $S(E)$ è equivalente alla entropia termodinamica ed in particolare che la temperatura del gas poteva essere calcolata:

$$\frac{1}{T} = \frac{\Delta S(E)}{\Delta E}$$

L'entropia di Boltzman

- A partire dalla formula:

$$S(E) = k_B \ln(\Omega(E))$$

Osserviamo:

- Se aumentiamo E come cambia il volume dei possibili stati delle particelle $\Omega(E)$?
- **Cresce se E cresce, decresce se E diminuisce.**
- Corrispondentemente come cambierà l'entropia $S(E)$?
- **Rivedersi le proprietà dell'esponenziale...**
- **Boltzman dimostrò che $S(E)$ è equivalente alla entropia termodinamica ed in particolare che la temperatura del gas poteva essere calcolata:**

$$\frac{1}{T} = \frac{\Delta S(E)}{\Delta E}$$

Osservazioni finali

- 1 ciccio...
- 2 tino...
- 3 pino...

Osservazioni finali

- 1 ciccio...
- 2 tino...
- 3 pino...

Osservazioni finali

- 1 ciccio...
- 2 tino...
- 3 pino...

Bibliografia

- *Caso, probabilità e complessità* A.Vulpiani, Ediesse (2015)
- *Caos* M. Malvaldi, S. Marmi, il Mulino (2019)
- *Information Theory, Inference and Learning Algorithms* D.Mackay, Cambridge UP (2008)