

# Il Concetto di Entropia

- L'entropia è un concetto legato al grado di “disordine” in un sistema.
- Il concetto è usato in vari settori delle scienze:
  - fisica (in particolare nella termodinamica)
  - teoria dell'informazione

# Teoria dell'Informazione (1)

- Entropia legata al concetto di “misura dell'informazione”
- Esperimento X1:
  - 4 possibili risultati: a, b, c, d equiprobabili
  - vogliamo memorizzare il risultato su un elaboratore: che codice utilizzare?

<i>Probabilità</i>	<i>Risultati</i>	<i>Codice binario</i>
$\frac{1}{4}$	a	00
$\frac{1}{4}$	b	01
$\frac{1}{4}$	c	10
$\frac{1}{4}$	d	11

- 2 bit per risultato

# Teoria dell'Informazione (2)

- Esperimento X2:
  - 4 risultati **non** equiprobabili

<i>Probabilità</i>	<i>Risultati</i>	<i>Codice binario</i>
$\frac{1}{2}$	a	0
$\frac{1}{4}$	b	10
$\frac{1}{8}$	c	110
$\frac{1}{8}$	d	1110

- codice con un numero di bit variabili
  - lunghezza media:  
 $\frac{1}{2} * 1 + \frac{1}{4} * 2 + \frac{1}{8} * 3 + \frac{1}{8} * 4 = \frac{15}{8}$
  - la lunghezza media è minore di 2!!
- questo codice per l'esperimento X1?
  - lunghezza media  $\frac{5}{2}$

# Teoria dell'Informazione (3)

- Perché per X1 sono necessari due bit mentre per X2 si può fare di meglio?
  - X2 contiene “meno informazione” di X1
- E` possibile fare ancora meglio per l'esperimento X2 ?
- A queste domande risponde la **teoria dell'informazione**.

# Entropia di un esperimento finito

- Sia  $X$  un esperimento con un numero finito di possibili risultati  $e_1, \dots, e_q$ .
- La probabilità che l'evento  $e_i$  si verifichi è  $p_i$ .
- Entropia dell'esperimento  $X$ :

$$H(X) = H(p_1, p_2, \dots, p_q) = - \sum_{i=1}^q p_i \log(p_i)$$

- Perché questa definizione?

# Proprietà della funzione H (1)

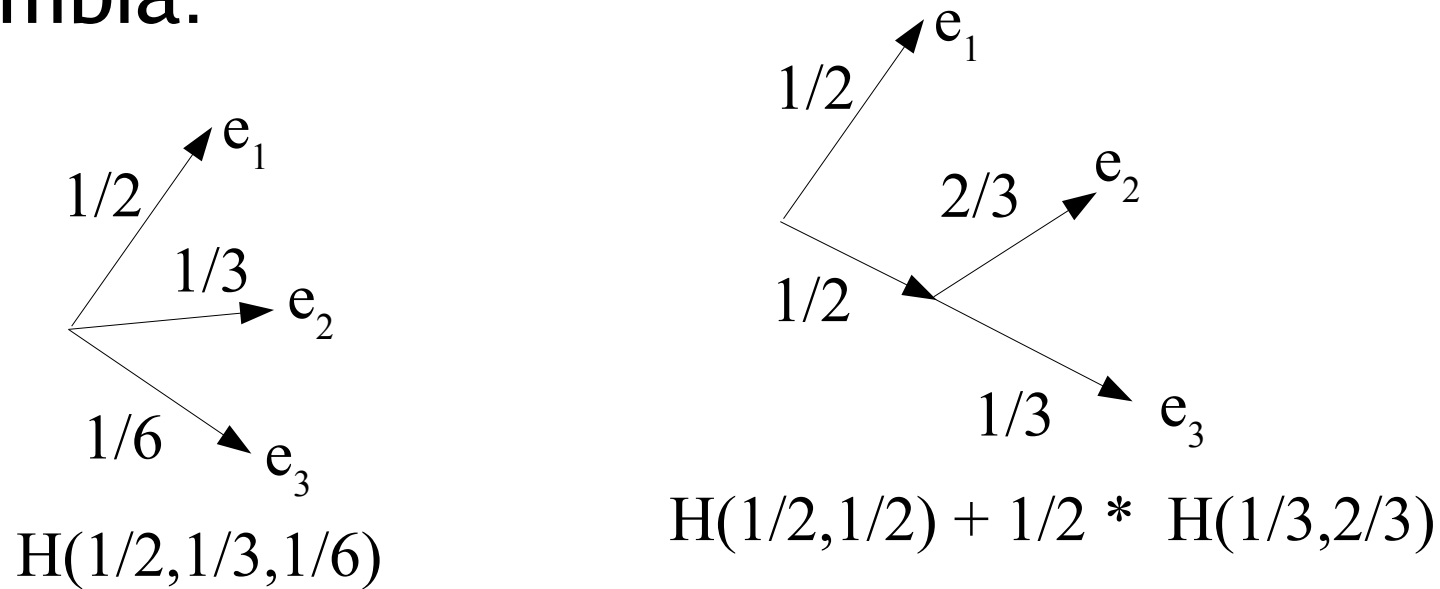
- H è continua sulle  $p_i$ 
  - piccole modifiche delle probabilità causano piccole modifiche della incertezza dell'esperimento
- se X e X' hanno q e q' risultati equiprobabili e  $q < q'$ , allora:

$$H(X) = H(1/q, \dots, 1/q) < H(1/q', \dots, 1/q') = H(X')$$

- più risultati possibili abbiamo maggiore è l'incertezza dell'esperimento

# Proprietà della funzione H (2)

- Se l'esperimento X è scomposto in due esperimenti successivi, il risultato non cambia:



– i due valori sono uguali

# Caratterizzazione di H

- Le uniche funzioni che soddisfano le proprietà di cui ai lucidi precedenti sono:

$$H_C(p_1, p_2, \dots, p_q) = -C \sum_{i=1}^q p_i \log(p_i)$$

al variare di C.

- Fissiamo  $C=1$  e chiamiamo **bit** l'unità di misura corrispondente a questa scelta.



# Entropia e codici

- Se  $X$  è un esperimento, un qualunque codice per  $X$  che sia unicamente decifrabile avrà una lunghezza media  $n \geq H(X)$ .
  - nell'esperimento  $X_1$  l'entropia è 2 per cui tutti i codici possibili avranno lunghezza media  $\geq 2$ .
  - nell'esperimento  $X_2$  l'entropia è  $14/8$ , il codice fornito ha lunghezza media  $15/8$ : forse si può migliorare
    - basta cambiare il codice del risultato  $C$  in  $111$ .

# Programmi di Compressione

- I programmi di compressione sfruttano il fatto che alcune combinazioni di caratteri sono più probabili di altre.
  - ovvero, l'esperimento “*prendo a caso un file di m bytes dal disco fisso*” ha una entropia inferiore ad  $8m$  (come sarebbe se il file fosse generato in modo completamente casuale)
  - vuol dire anche, in ogni sistema di compressione, ci saranno dei file in cui la versione compressa è più grande di quella originaria.