Il caso e la necessità

Dio non gioca a dadi con l'universo, Egli è sottile ma non malizioso (A. Einstein)

Causalità, casualità, e pseudo-casualità

Causalità

Il gioco del biliardo è un gioco di precisione. Conoscendo esattamente l'intensità e la direzione del tiro, il movimento di tutte le palle potrebbe essere calcolato utilizzando le leggi della dinamica (il condizionale è da intendere "a meno di errori di misura e di calcolo").

Casualità

Il gioco della roulette è un gioco di azzardo puro. A gioco corretto tutte le precauzioni vengono prese affinché il risultato di ogni lancio sia completamente imprevedibile e la probabilità di uscita di ognuno dei 37 numeri sia esattamente 1/37.

> Entrambi gli strumenti sono meccanici, obbediscono alle stesse leggi fisiche, e in entrambi vi è l'intervento manuale di un operatore (il giocatore di biliardo o il croupier).

La differenza è che nel biliardo tutto è fatto per minimizzare gli effetti del caso (superficie piana e perfettamente orizzontale, panno liscio, palle rotonde, pesanti e pulite). Nella roulette invece tutto massimizza gli effetti del caso (il piatto ruota e le losanghe servono a far rimbalzare la leggera pallina in modo imprevedibile).

Nella realtà, al livello macroscopico che cade sotto i nostri sensi, il caso è spesso un velo che ci impedisce una conoscenza deterministica.

Nella fisica classica gli approcci statistici sono una scorciatoia potentissima adottata quando le variabili in gioco sono troppe e non si conoscono le condizioni iniziali.

In questo senso una sequenza di numeri generati da una macchina (tipo quelle per le lotterie) in via teorica potrebbe essere prevista ma in pratica è completamente imprevedibile ed è considerata casuale.

Nella meccanica quantistica l'approccio probabilistico è inevitabile e ogni tentativo di farne a meno sembra essere finora fallito. Era questo che dava fastidio ad Albert Einstein che la considerò fino all'ultimo una teoria "incompleta".

Una sequenza di numeri generata da un rivelatore di fenomeni microscopici (per esempio emissioni radioattive) è casuale anche da un punto di vista teorico.

Pseudo-casualità

Una sequenza di numeri è pseudo-casuale se

- è facilmente generabile con un programma;
- gode delle stesse proprietà statistiche di una sequenza casuale.

Esempio

Questi sono veri numeri del lotto.

```
29 52 31 80  7 18 44 29 50  1 29 21 46 84 13 87 48 28 78  4 32 34 24 83 41  1 81  2 36 68 13 62 44 55 12 56 65
2  5 81 64 14 54 83 53 81 43 87 37 45 18 50 85  3 42 45 15 25 74 23 90 59 46 30 52 51  9 89 54 36 34 19 60 84
17 89 58 84 11 32 15 30 83 82  3 12 89 54 66 86 58  1 21 19 59  2 39 70 72 66 12 62 51 88 30 83 84 38 78 59 70
15 50 59 10 64  7 53 30 85 73 52  3 89 76 86 48 53 57 18  2 25 13 74 36 63 50  1 26 75 5 66 17 22 21 60 81 87
42 24 43 79 15 38 90 16 49  6  8 86 58 27 13 33  1 53  9 68  5 61 47  4 30 16 27 86 89 47  8 73 22 40 74 39 77
88 77 85 67 18 60 77 63 79  5 72  4 68 18 50 ...
```

Questi sono numeri pseudo-casuali generati al computer.

```
74 75  2 27 28  2  7 60 73 41  3 60 22 41 68 19 64 47 13  2 76  3 41 49 63 85 73 65 37 42 23 27  3 55 36 46 87
29 78 90 52 65 82 30 58 65 22  4 84 89  2 10  6 28 26 69 24 37 68 56 63 32 37 22 57 45 81 42 86 70 46 55 35 56
75 65 82 68 30 37 81 39 63 46 71 42 53 44 90 72 39  7 75 26 16 39 32 49 12 37 46 50 53 85 75 22 83 48 90 47 43
12 33 62 16 69 82 14 58 22 57 85 59 38  8 72 20  4 73 60 30 26 29 39 62 78 16 63 13 59 62 21 47 58 17 20 86  3
65 69 22 44 13 75 48 61 36 10 50 10 53 31  7 64 90  3 43 34 10 31 82 79 24 29 82 36 15  1 57 65 32 50 14 84 81
84 63 43 26 42 12 87 57 26 53 19 43 89 18 59 ...
```

A occhio le due sequenze sono indistinguibili, la differenza sta nel fatto che non esiste alcun programma che possa generare la prima non se quello che esegue la banale enumerazione.

Con la terminologia della complessità program-size la prima sequenza è casuale e la seconda no, ma, a posteriori, in base al Teorema di Chaitin, se le sequenze sono abbastanza lunghe esse sono **algoritmicamente indistinguibili**.

Il Lotto (*la madre dei fessi è sempre incinta*)

Il gioco del Lotto e le lotterie sono stati introdotti per migliorare le finanze statali (non solo in Italia e non solo in tempi recenti, uno dei primi fu Casanova, per il Re di Francia). Una semplice analisi del regolamento mostra che il Lotto è un gioco non equilibrato in cui il banco (lo Stato) guadagna sempre. Anche la Roulette è un gioco squilibrato (per la presenza dello zero) ma in confronto il Lotto è **molto** più svantaggioso.

> L'unico modo razionale (sic!) per giocare al Lotto è aspettare che la buonanima di un parente defunto ci appaia in sogno e ci dia qualche numero buono.

E' abbastanza comune la credenza che lo studio dei risultati passati possa portare qualche informazione sui risultati futuri e in particolare che una strategia di gioco basata sui ritardi abbia qualche speranza di guadagno.

> Purtroppo una semplice analisi della tecnologia fisica con cui sono effettuate le estrazioni mostra senza ombra di dubbio che, imbrogli a parte, le estrazioni sono indipendenti e che quindi **in nessun modo** il passato può influenzare il futuro.

Tutte le strategie di gioco hanno la **stessa** speranza di vincita qualunque sia il numero giocato (sia se si giocano sempre gli stessi numeri sia se si cambia ogni volta). Questo è banalmente ovvio, per chi è in grado di ragionare, se si pensa che i numeri sono solo nomi accidentali e che il gioco è invariante per permutazione.

In altre parole le proprietà statistiche del gioco non dipendono (**né potrebbero dipendere in alcun modo**) dai numeri scritti sui foglietti all'interno delle palle.

Chi usa la Legge dei Grandi Numeri per giustificare la teoria dei ritardi è un **idiota ignorante** oppure un **truffatore**.

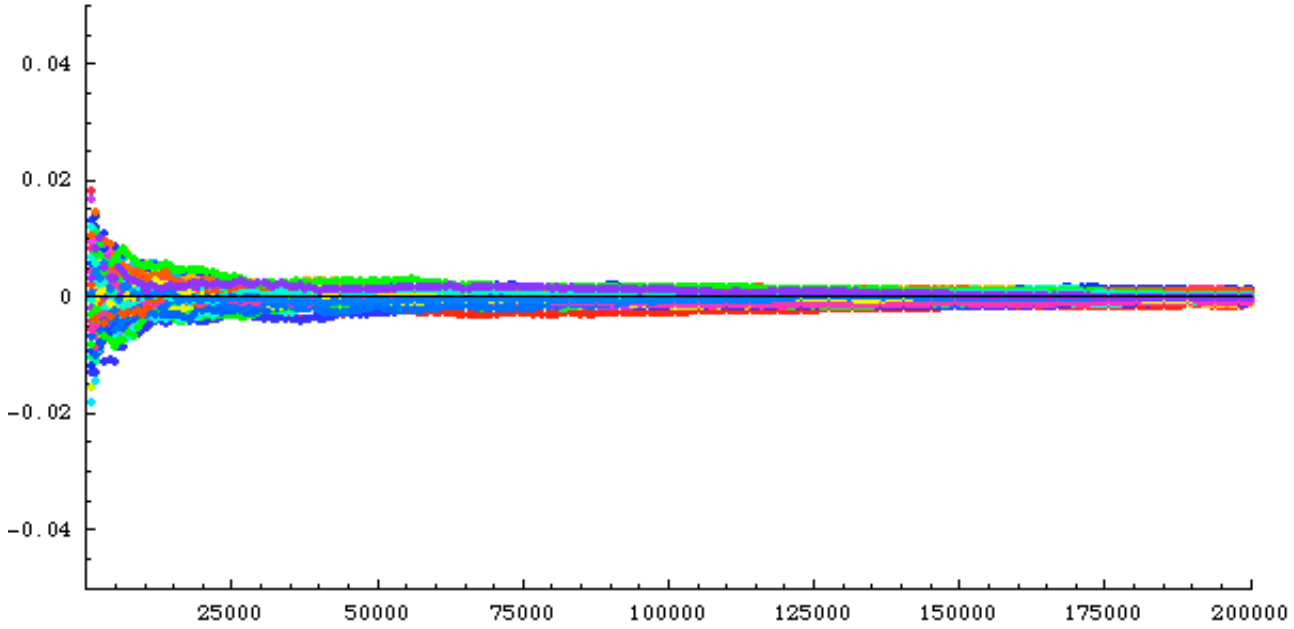
La legge dei grandi numeri

La frequenza media tende al valore aspettato.

Per qualunque $\varepsilon>0$ prefissato piccolo a piacere,

$$\lim_{n \to \infty} \Pr.\left(\left| \frac{S_n}{n} - \mu \right| > \varepsilon \right) = 0$$

Applicato al Lotto questo significa che il rapporto tra le uscite e le estrazioni di ogni numero tende a 5/90 = 1/18 quando il numero delle estrazioni tende all'infinito.



Differenza tra il rapporto uscite/estrazioni e la media teorica 1/18 per i 90 numeri del lotto su 200000 estrazioni. Secondo la LDGN, il valore tende a zero al crescere di n.

La LDGN **non** afferma assolutamente che lo scarto $S_n - \mu n$ tenda a zero per n che tende all'infinito. Per avere qualche informazione sul comportamento dello scarto bisogna scomodare il Central Limit Theorem.

Central Limit Theorem (Zentraler Grenzwertsatz), *Teorema fondamentale sui limiti*.

Per ogni x reale vale:

$$\lim_{n \to \infty} \Pr\left(S_n - n\mu \leq x\sigma\sqrt{n}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$

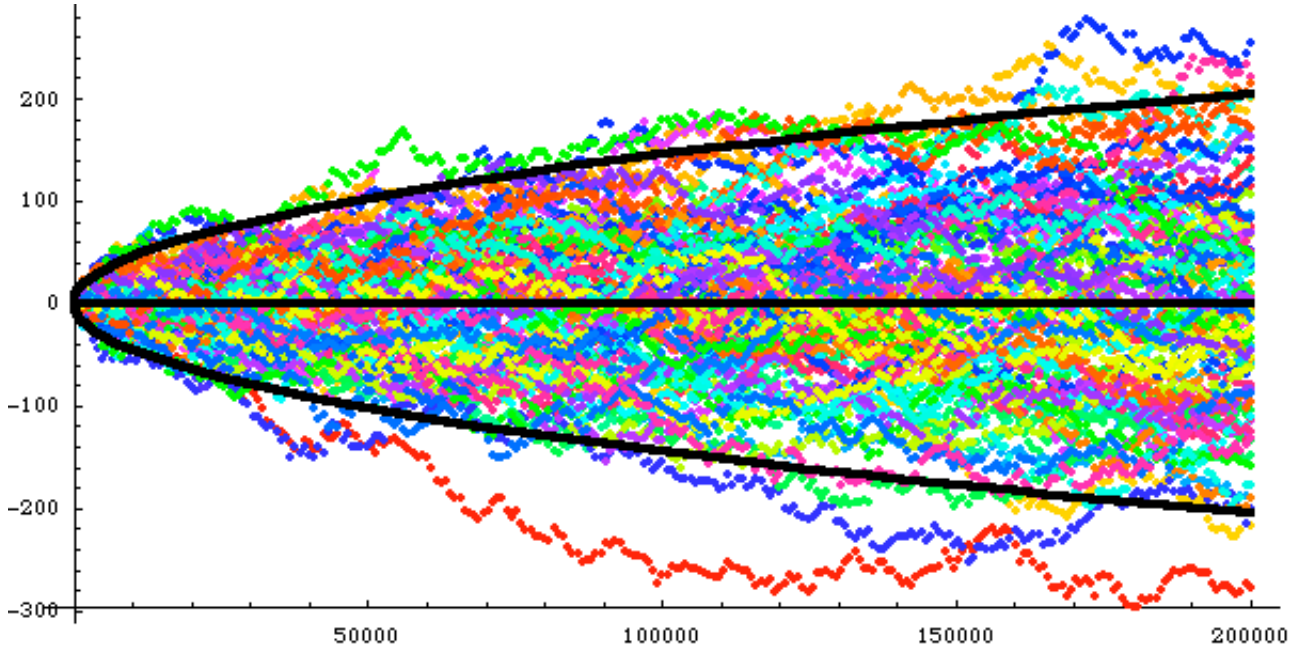
Il termine al secondo membro vale praticamente 1 per x>3. Quindi, il numero di uscite effettivo si avvicina al numero di uscite teorico con un errore che va come la radice di n.

In altre parole (e meno rigorosamente) si può dire che

$$\frac{S_n - n\mu}{\sigma\sqrt{n}}$$

tende ad una gaussiana di media 0 e varianza 1.

In base al CLT lo scarto può andare all'infinito (SORPRESA!!) ma non più velocemente della radice di n.



Differenza tra il numero delle uscite e il valore atteso n/18 per i 90 numeri del lotto su 200000 estrazioni. Secondo il CLT può crescere come la radice di n.

Ed ecco distrutte in un colpo solo tutte le farneticazioni dei ritardisti.

Entropia e bit

Entropia termodinamica

Uno dei concetti fondamentali della fisica (e dello scibile umano): l'entropia rappresenta una misura del disordine di un sistema.

> Se un sistema riceve ua quantità di calore Δq ad una temperatura assoluta T questo causa un aumento $\Delta S = \Delta q/T$ dell'entropia del sistema.

Il secondo principio della termodinamica: in un sistema isolato l'entropia non può diminuire.

Corollario: se in processo l'entropia termodinamica aumenta il processo è irreversibile.

Esempi.

- passaggio di calore da un corpo caldo ad uno freddo;
- trasformazione di moto in calore (pala di ventilatore spento).
- *se fai bollire un acquario ottieni una zuppa di pesce, ma è molto difficile che raffreddando la zuppa di pesce ritorni ad avere l'acquario*.

Formulazione statistica dell'entropia. In un sistema le cui componenti hanno distribuzione di energia w_n vale

$$S = -\sum_n w_n \log w_n$$

NB. vi sono molte varianti e molti modi per formulare la definizione statistica che vale anche a livello di fisica quantistica.

Entropia informatica

L'incertezza nel risultato di un esperimento con probabilità p_i vale

$$H = -\sum_i p_i \log_2 p_i$$

si misura in bit e viene detta entropia dell'esperimento. La formula si può derivare da alcuni semplici e ragionevoli assiomi sul contenuto informativo di un esperimento stocastico. Il nome deriva dal fatto che Shannon sapeva la fisica.

Relazioni tra l'entropia termodinamica e informazione

> Ogni esperimento che porta ad un aumento dell'informazione costa un aumento maggiore di entropia termodinamica (Brillouin).
>
> Ogni conoscenza che può portare ordine si paga con un disordine fisico non minore.

Entropia ed evoluzione

Il secondo principio della termodinamica spesso viene interpretato come una tendenza naturale obbligatoria al disordine. In realtà all'interno di sistemi aperti o in parti di sistemi chiusi lontane dall'equilibrio nulla impedisce il verificarsi di fenomeni di auto-organizzazione. Chi vi parla è uno di questi (e non nel senso che sono bravo a tenere le penne in ordine).

Nei casi in cui si può fare una analisi precisa si vede che la organizzazione locale avviene sempre a spese dell'entropia termodinamica globale.

Esempi di auto-organizzazione

Fisica

- cristallizzazione
- laser
- superconduttori
- turbolenza, convezione
- formazione di strutture in astrofisica (stelle, galassie)
- terremoti, valanghe, incendi
- ingorghi, blackout

Chimica

- reazioni oscillanti
- reti autocatalitiche

Biologia

- omeostasi (capacità di un sistema di mantenere un equilibrio interno stabile anche in presenza di variazioni delle condizioni esterne).
- strutture sociali (formiche, api)
- origine della vita
- evoluzione
- morfogenesi

Matematica e Informatica

- automi cellulari
- calcolo evolutivo
- intelligenza artificiale

Sociologia

- comportamento sociale organizzato (spesso con proprietà statistiche tipo legge di Zipf)

L’evoluzione è un fatto (come la terra rotonda)

L’evoluzione biologica darwiniana aumenta l’ordine a spese dell’entropia termodinamica (aumento locale - perdita globale). Si dice che l’evoluzione biologica si mantiene al margine del caos.

Teoria del gene egoista. Secondo Dawkins il motore dell’evoluzione non è la sopravvivenza degli individui, dei gruppi o delle specie ma di quegli insiemi di informazione che (semplificando) egli chiama geni e che possono durare milioni di anni. Nessuna visione teleologica ma solo un dato di fatto matematico.

I replicanti capaci di sopravvivere meglio e più a lungo, replicano più degli altri!

In realtà quella che si propaga e sopravvive è l’informazione (versione biologica del vivo nelle mie opere). I sistemi auto-organizzati hanno invece una durata limitata nel tempo.

Ovvero: Memento homo qui pulvis es, et in pulverem reverteris

(in questo contesto da intendere in senso termodinamico e non religioso)

La freccia del tempo

Che cosa è dunque il tempo? Se nessuno me ne chiede, lo so bene: ma se volessi darne spiegazione a chi me ne chiede, non lo so: così posso dire di sapere che se nulla passasse, non vi sarebbe il tempo passato, e se nulla sopraggiungesse, non vi sarebbe il tempo futuro, e se nulla fosse, non vi sarebbe il tempo presente. Ma in quanto ai due tempi passato e futuro, in qual modo essi sono, quando il passato, da una parte, più non è, e il futuro, dall'altra, ancora non è? In quanto poi al presente, se sempre fosse presente, e non trascorresse nel passato, non più sarebbe tempo, ma sarebbe, anzi, eternità. Se, per conseguenza, il presente per essere tempo, in tanto vi riesce, in quanto trascorre nel passato, in qual modo possiamo dire che esso sia, se per esso la vera causa di essere è solo in quanto più non sarà, tanto che, in realtà, una sola vera ragione vi è per dire che il tempo è, se non in quanto tende a non essere? (Agostino, Le confessioni)

Il tempo è una cosa strana. Passiamo così i giorni della vita e nulla è il tempo. Ma poi ad un tratto, ecco, altro non sentiamo che lui. E' intorno a noi, è anche dentro di noi. Sui volti cola, cola nello specchio e scorre nelle mie tempie. Silente, come una clessidra. Talvolta io l'odo che scorre senza sosta. (H. von Hofmannsthal, monologo della Marescialla, dal Cavaliere della Rosa)

Istante fermati, sei bello! (Goethe, Faust)

Il tempo è sinonimo di irreversibilità:

Irreversibilità termodinamica. Ogni processo che causa aumento di entropia è irreversibile. Nessuno ha mai visto una ruota che si mette in moto da sola perché si sta raffreddando.

Irreversibilità per perdita di informazione. Ogni qual volta dell'informazione viene distrutta il processo è irreversibile. Per i legami tra le due entropie la distruzione di informazione è un processo che costa energia e produce aumento di entropia termodinamica.

Irreversibilità per l'osservazione microscopica, In meccanica quantistica l'osservazione altera permanentemente lo stato del sistema (il gatto di Schroedinger si decide a essere tutto vivo o tutto morto).

Ιόπλοκ' ἄγνα μελλιχόμειδε Σάπφοι

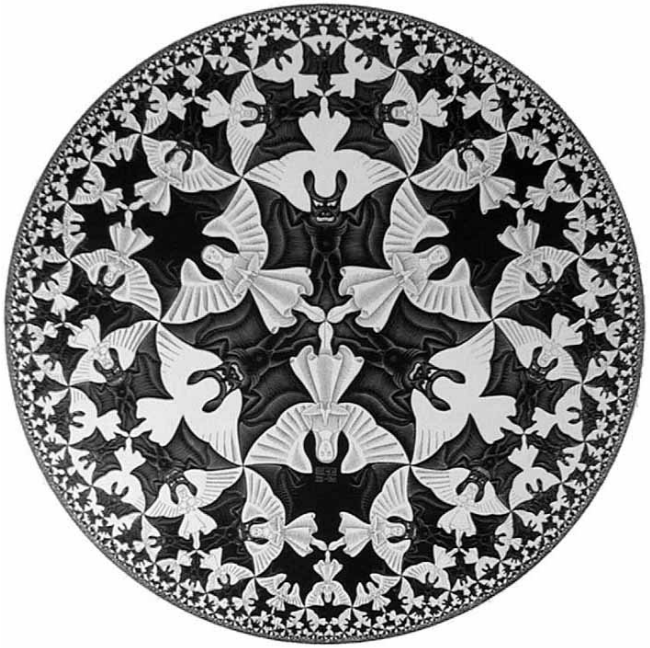
irraggiungibile nello spazio e nel tempo come la stella più lontana,

Per essere irraggiungibili nello spazio e nel tempo

non servono 26 secoli,

possono bastare 10 secondi.

Informazione e Ridondanza



Conticuere omnes intentique ora tenebant.

inde toro pater Aeneas sic orsus ab alto:

infandum, regina, iubes renovare dolorem

(Virgilio)

Cosa è la Ridondanza

La ridondanza si può definire informalmente come un eccesso di informazione: qualcosa che si può togliere senza perdere i contenuti fondamentali che si intende comunicare.

È importante sottolineare che la ridondanza NON è mera ripetizione anche se spesso esistono trasformazioni dei dati che possono evidenziare le ripetizioni.

In genere anche la presenza di una struttura implica una ridondanza.

Prima di discutere questo concetto in modo più rigoroso vediamo qualche esempio:

- nelle lingue
- nell'audio
- nelle immagini

Ridondanza nelle lingue

E' facile riconoscere la lingua che si sta ascoltando spesso anche senza comprendere ciò che vien detto.

Ridondanza senza significato

Generazione di frasi casuali che soddisfano alcune regole sintattiche.

Gli esempi sono tratti dal testo, ormai classico, Abramson, Information Theory and Coding, McGraw-Hill, 1963.

Approssimazioni di ordine 3

Testi casuali che rispettano la frequenza delle triple di varie lingue.

Ianks can ou ang rler thatted of to shor of to havem a i mand and but whissitably thervereer eights takillis ta

Jou mouplas de monnernaissains dem us vreh bre tu de toucheur dimere ll es mar balme re a ver douvents so

Bet ereiner sommeit sinach gan turhatt er aum wie best alliebder taussichelle laufurcht er bleindeseit uber konn

rame de lla el guia imo sus condias su e uncondadado dea mare to buerbali a nuae y hararsin de se sus suparoceda

Et ligercum siteci libemus acererlen te vicaescerum pe non sun minus uterne ut in ario popomin se inqueneque ira

Approssimazioni di ordine 3

Testi casuali che rispettano la frequenza delle triple di varie lingue.

Ianks can ou ang rler thatted of to shor of to havem a i mand and but whissitably thervereer eights takillis ta INGLESE

Jou mouplas de monnernaissains dem us vreh bre tu de toucheur dimere ll es mar balme re a ver douvents so FRANCESE

Bet ereiner sommeit sinach gan turhatt er aum wie best alliebder taussichelle laufurcht er bleindeseit uber konn TEDESCO

rame de lla el guia imo sus condias su e uncondadado dea mare to buerbali a nuae y hararsin de se sus suparoceda SPAGNOLO

Et ligercum siteci libemus acererlen te vicaescerum pe non sun minus uterne ut in ario popomin se inqueneque ira LATINO

Ed ecco una approssimazione dell'INGLESE di ordine 5

The head and in frontal attack on an english writer that the character of this point is therefore another method for the letters that the time of who ever told the problem for an unexpected

Uso della Ridondanza Linguistica

- La ridondanza lessicale e grammaticale porta con sé una buona capacità di correzione che di per sé migliora le possibilità di comprensione.
  - lettura veloce
  - comprendere un parlato in un ambiente rumoroso;
  - decifrare un testo in parte corrotto
  - acquisire le abilità che permettono di scrivere sotto dettatura

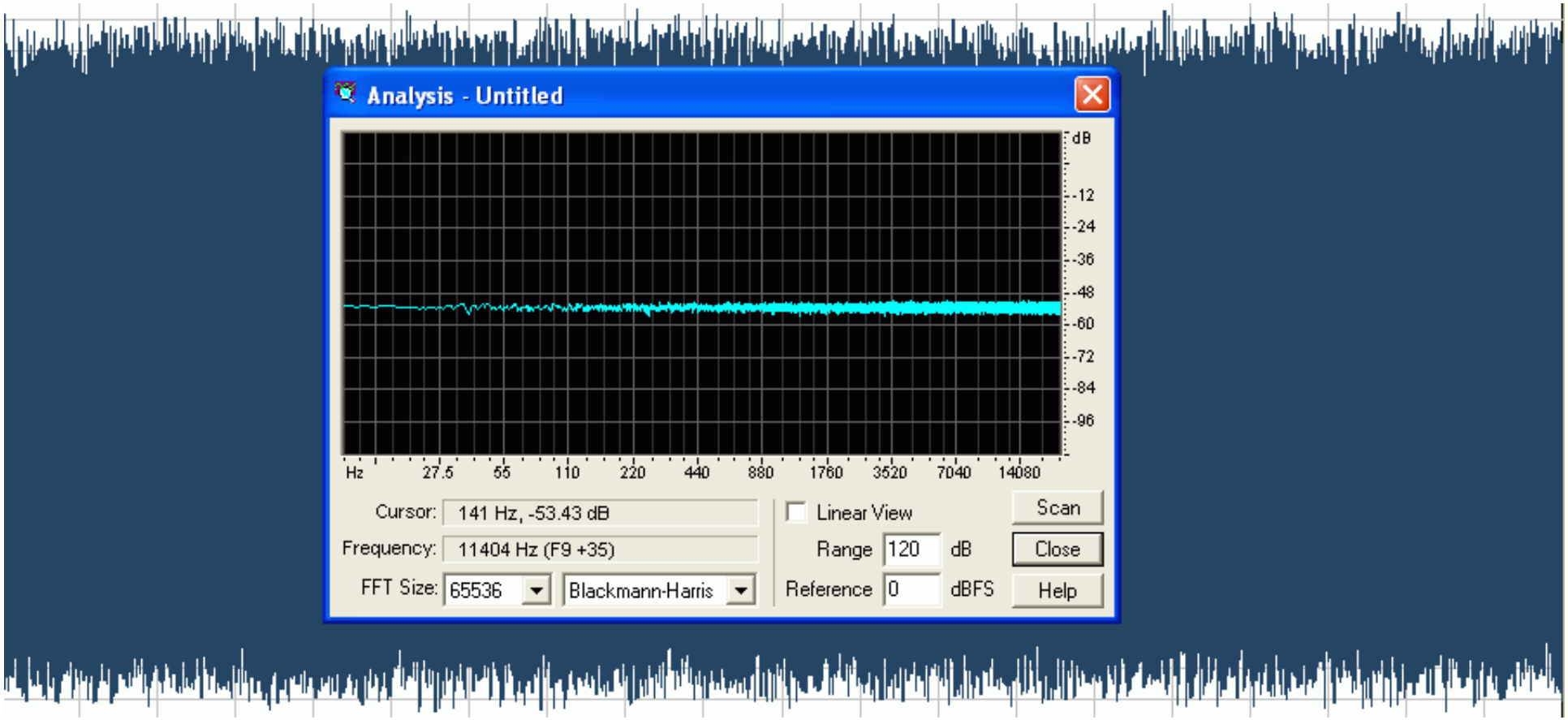
- Talvolta si codificano le singole lettere per permettere una agevole comprensione di parole non banali o che devono essere trasmesse senza errori, introducendo una ulteriore ridondanza.
  - codice: Alfa, Bravo, Charlie, Delta ... Victor, Whiskey, Xray, Yankee, Zulu
  - codice: nomi di città

- Compressione
  - riassunto
  - messaggini

Esempi Audio

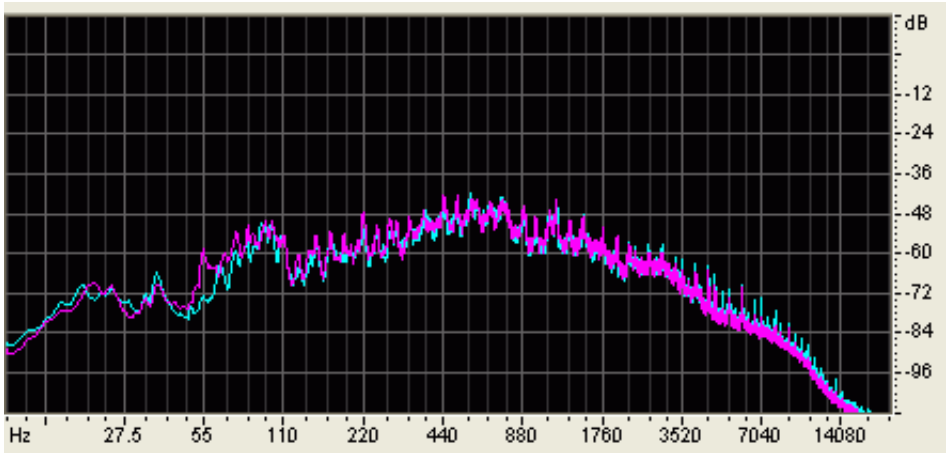
Un suono assolutamente non ridondante è il rumore bianco. Il vero rumore bianco ha solo una valenza teorica, porta una informazione infinita e non è generabile in natura.

Un segnale digitalizzato in cui ogni campione è indipendente da tutti gli altri costituisce uan approssimazione udibile al rumore bianco e porta la MASSIMA informazione possibile in quel formato.



Da un punto di vista artistico il rumore bianco fa schifo.

Audio ridondante



Bach BWV 108, 4 - coro

A rigorous piece of choral polyphony, three tersely arranged fugues in motet style (J.E. Gardiner)
Le fughe sono a 4 voci (Basso, Tenore, Alto, Soprano) ogni versetto viene ripetuto 8 volte il tutto in 2' 35"

Wenn aber jener, der Geist der Wahrheit, kommen wird, der wird euch in alle Wahrheit leiten.
Denn er wird nicht von ihm selber reden, sondern was er hören wird, das wird er reden;
und was zukünftig ist, wird er verkündigen.

Ma quando Lui, lo Spirito di verità,verrà, vi guiderà alla verità tutta intera.
Egli non parlerà da sè stesso, ma tutto quello che ascolterà, ve lo dirà;
e vi annuncerà le cose future (Giovanni 16,13)

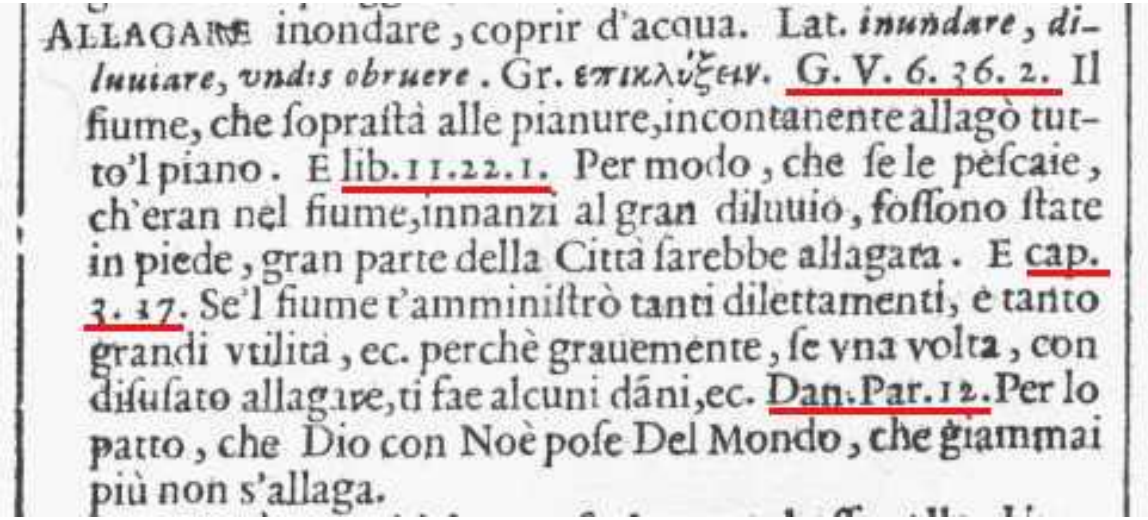
Codifica a blocchi

Dato un alfabeto $S=\{s_1, s_2, \ldots, s_k\}$ e un alfabeto $G=\{a_1, a_2, \ldots, a_r\}$,

si dice codice a blocchi una funzione che associa ad ogni elemento di S (il messaggio) una e una sola sequenza di simboli di G. Le sequenze di simboli di G si chiamano parole del codice.

Esempi storici

- La scrittura sillabica o alfabetica nelle sua varie forme
- Le abbreviazioni bibliografiche o in epigrafi (esempio dal Vocabolario della Crusca)



ALLAGARE inondare, coprir d'acqua. Lat. *inundare, diluuiare, undis obruere*. Gr. επικλύζειν. G. V. 6. 36. 2. Il fiume, che sopraſtà alle pianure, incontanente allagò tutto'l piano. E lib. 11. 22. 1. Per modo, che ſe le peſcaie, ch'eran nel fiume, innanzi al gran diluuio, foſſono ſtate in piede, gran parte della Città ſarebbe allagata. E cap. 3. 17. Se'l fiume t'amminiſtrò tanti dilettamenti, e tanto grandi vtilità, ec. perchè grauemente, ſe vna volta, con diſuſato allagare, ti fae alcuni dãni, ec. Dan. Par. 12. Per lo patto, che Dio con Noè poſe Del Mondo, che giammai più non s'allaga.

Codifica per crittografare

Messaggi segreti (un'idea vecchia come il mondo)

m messaggio

k chiave

c messaggio cifrato

Codifica

c = F(m,k) F deve essere facile da calcolare

Decodifica

m = G(c,k) G(.,k) deve essere facile da calcolare

Se non si conosce k allora G deve essere impossibile o difficilissima da calcolare

Nello “Scarabeo d’oro” di Edgar Allan Poe vi è un bell’esempio di decodifica basata su tecniche statistiche.

“53‡‡†305))6*;4826)4‡.)4‡);806*;48†8
¶60))85;1‡(;:‡*8†83(88)5*†;46(;88*96
?;8)‡(;485);5*†2:*‡(;4956*2(5*—4)8
¶8*;4069285);)6†8)4‡‡;1(‡9;48081;8:8‡
1;48†85;4)485†528806*81(‡9;48;(88;4
(‡?34;48)4‡;161;:188;‡?;”

Of the character 8 there are 33.
; “ “ 26.
4 “ “ 19.
‡) “ “ 16.
* “ “ 13.
5 “ “ 12.
6 “ “ 11.
†1 “ “ 8.
0 “ “ 6.
92 “ “ 5.
:3 “ “ 4.
? “ “ 3.
¶ “ “ 2.
—. “ “ 1.

5 represents a
† ” d
8 ” e
3 ” g
4 ” h
6 ” i
* ” n
‡ ” o
(” r
; ” t
? ” u

Il risultato è

A good glass in the bishop’s hostel in the devil’s seat forty-one degrees and thirteen minutes northeast and by north main branch seventh limb east side shoot from the left eye of the death’s-head a bee-line from the tree through the shot fifty feet out.

Compressione

I primi esempi di compressione dei dati risalgono agli albori della scrittura, quando uno scalpellino si trovò a combattere con una lastra di pietra troppo piccola e un testo troppo lungo e fu costretto ad abbreviare le parole

Un esempio più recente, ma sempre stagionato, è rappresentato dal Vocabolario degli Accademici della Crusca che nella edizione del 1612 presenta un'intera serie di abbreviazioni bibliografiche tutte diverse tra loro usate con lo scopo di far tornare le lunghezze delle righe (pazientemente composte a mano).

È però con l'avvento della memorizzazione elettronica dei dati che il problema del supporto piccolo e dei dati ingombranti si propone in tutta la sua drammaticità e la vendita di compressori diviene un ottimo affare.

Vi è sempre un limite alla compressione (contenuto informativo), inoltre

Le stringhe corte sono poche, quelle lunghe sono tante!!!

A seconda della utilizzazione si distinguono tra tecniche di memorizzazione compressa (i ben noti "Zippatori") e tecniche di trasmissione compressa (ovvero le tecniche usate nei Modem per aumentare l'efficienza di trasmissione a parità di costo). Ma la vera distinzione è da fare tra compressione *lossless* (priva di perdite) e compressione *lossy* (con perdita di informazione).

- Il primo tipo di compressione, che deve permettere la ricostruzione senza errori dei dati originali, si basa soprattutto sulla eliminazione delle ridondanze e per questo presenta limiti ben precisi. L'uso di questo tipo di compressione è indispensabile per memorizzare o trasmettere programmi, testi e ogni tipo di informazione che non può essere alterata senza danni.

- Nel secondo caso si accetta di perdere informazione avendo in cambio il vantaggio di non avere limiti al tasso di compressione. Questa codifica si presta bene ad essere usata per dati come suoni e immagini che per la loro natura sono comunque soggetti ad una inevitabile perdita di informazione (causata dal rumore ambientale, la risoluzione dei trasduttori, ecc.)

Compressione delle immagini

Nel caso delle immagini vi sono due formati classici che incarnano le due tecniche, il formato **GIF**, privo di perdite e adatto alle immagini piccole, e il formato **JPEG**, capace di ridurre maggiormente le dimensioni, più adatto a conservare immagini di grande formato introducendo però una degradazione della qualità.

Ridondanza come compressibilità

La Teoria dell'Informazione di Shannon fornisce un quadro sistematico in cui i concetti di ridondanza e compressione hanno una collocazione rigorosa.

Teorema della codifica in assenza di rumore

Si può dimostrare che data una sorgente di informazione S la sua entropia H(S) è il limite inferiore al numero di bit necessari per trasmettere i dati di S senza perdite.

Dopo una compressione ottima nel senso di Shannon i bit sono indipendenti ed equiprobabili ovvero è stata eliminata ogni ridondanza.

Ridondanza per correggere

Se si codifica una sorgente togliendo ogni ridondanza basta perdere un bit per non poter più ricostruire i dati originali.

Canale di trasmissione: un modello probabilistico di come i bit possono venire alterati nel processo di tramissione.

Capacità del canale

$$C = \max_{E} H(E) - H(E/U)$$

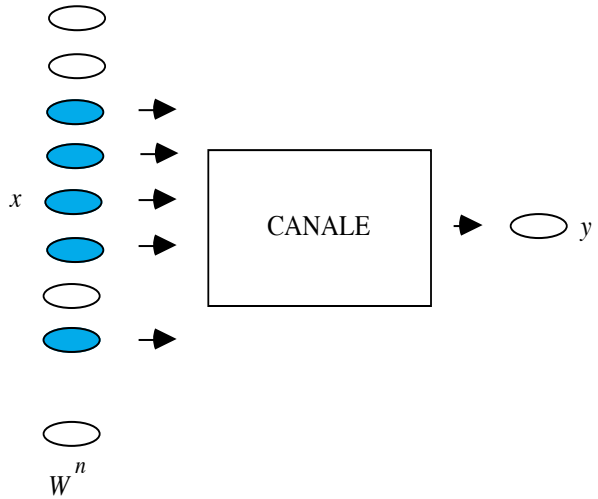
Per ovviare alla possibilta di perdere bit si introduce una ridondanza che permetta di ricostruire i dati originali anche in presenza di (alcuni) bit alterati.

Può sembrare una contraddizione ma in realtà quello che si fa è

- Togliere la ridondanza caratteristica della sorgente, con un compressore
- Aggiungere una ridondanza adatta al canale di trasmissione, con un codice correttore

Teorema del panettone con l'uvetta.

Sia dato un canale con capacità C. Per qualunque $R < C$ è possibile trovare una sequenza di codici di lunghezza crescente n con velocità di trasmissione R e probabilità di errore che tende a 0.



Conclusioni

Un oggetto strutturato

- può essere compresso (perdendo struttura)
- oppure arricchito per aumentarne la robustezza

L'assenza di ridondanza porta la massima informazione e la massima casualità ma è priva di bellezza.

L'arte è struttura e la struttura arte