

Fondamenti di Informatica

Un computer è detto anche calcolatore, elaboratore oppure ordinatore ed è un dispositivo fisico che implementa il funzionamento di una macchina di TURING.

Possiedono tutti 4 cose:

- Un'unità centrale di elaborazione (CPU)
- Una memoria
- Un dispositivo d'ingresso/uscita
- Un canale perché possano dialogare tra loro, in accordo all'architettura di von Neumann

Alan Turing introdusse la macchina ideale ed il test che portano il suo nome.

Un calcolatore elettronico che nasce per eseguire programmi; il programma più importante per un CE è il sistema operativo che si occupa di gestire la macchina, le sue risorse e le altre applicazioni che vi sono eseguite. Fornisce all'utente un'interfaccia per inserire ed eseguire gli altri programmi. Questa parte immateriale vengono chiamate **software** in contrapposizione all'**hardware** che è la parte fisica degli elaboratori.

La macchina di **VON NEUMANN** è uno schema di progettazione di calcolatori elettronici che prende il nome dal matematico in questione che fu sviluppato per il sistema IAS di macchine.

Lo schema di base su cinque componenti fondamentali:

- CPU
- UNITA' DI MEMORIA intesa come memoria di lavoro
- UNITA' DI INPUT tramite la quale i dati vengono inseriti nel calcolatore per l'elaborazione
- UNITA' DI OUTPUT necessaria affinché i dati elaborati possano esser restituiti all'operatore
- BUS , un canale che collega tutti i componenti tra loro

Vannevar Bush : memex (memory expansion)

Ted Nelson : hypertext e interfaccia utente semplice

Moore: prima e seconda legge di Moore sulla complessità di un microcircuito.

Doug Englebart : primo uso del mouse nel 1968

Tim Berners Lee : world wide web (1989) HTML E HTTP (1989-1990)

Larry Page: google (1998)

Mark Zuckerberg: Facebook 2004

La **PIRAMIDE DIKW** detta anche **gerarchia dell'informazione** definisce un modello di rappresentazione delle relazioni tra **saggezza-esperienza, conoscenza, informazioni e dati**.

Un **dato** è una rappresentazione codificata da un segno/simbolo/segnale di un'entità ed i computer memorizzano ed elaborano dati, che rappresentano numeri, lettere, immagini, suoni, video ai quali viene attribuito un **significato**, affinché rappresentino una porzione di realtà.

Un'informazione è un insieme di uno più dati che si possiede e che serve ad acquisire una conoscenza. Un'informazione è sempre legata ad un codice che serve a rappresentarla, ad un

elemento che la contiene come un messaggio ed un mezzo attraverso il quale può essere trasmessa (media). L'informazione può essere: **analogica o digitale**.

La maggior parte delle grandezze fisiche sono di tipo continuo. I dati in forma **analogica sono valori di una grandezza fisica rappresenta in un insieme continuo; i dati in forma digitale sono rappresentati in un insieme di dimensioni finite**.

I dati si possono suddividere in diverse categorie o tipi:

- **Dati semplici come un numero, un carattere , una parola o un segno**
- **Dati complessi come una enumerazione di più dati semplici o complessi organizzati in struttura.**

Tipicamente l'informazione è definita in termini di dati e la conoscenza in termini di informazioni scambiate attraverso media.

L'unità di misura dell'informazione è il **bit, definito come la quantità minima di informazione che serve a discernere tra due possibili alternative ed è una cifra BINARIA, ovvero uno dei due simboli del sistema numerico binario, classicamente detti zero "0" e uno "1"**.

Come codifica binaria, il Bit rappresenta l'unità di definizione di uno stato logico o unità elementare dell'informazione trattata da un elaboratore. Un modo conveniente di raggruppare i bit è il **BYTE, composto da una sequenza di 8 bit**.

Siccome il byte è formato da 8 bit, esso è in grado di assumere $2^8 = 256$ possibili stati (da 0 a 255).

Multipli di bit e byte (standard SI)

NOME	Simbolo	multiplo	NOME	Simbolo	multiplo
bit	bit		byte	B	1B = 8 bit
kilobit	kbit	10^3 (~ 2^{10})	kilobyte	kB	10^3 B
megabit	Mbit	10^6 (~ 2^{20})	megabyte	MB	10^6 B
gigabit	Gbit	10^9 (~ 2^{30})	gigabyte	GB	10^9 B
terabit	Tbit	10^{12} (~ 2^{40})	terabyte	TB	10^{12} B
petabit	Pbit	10^{15} (~ 2^{50})	petabyte	PB	10^{15} B
exabit	Ebit	10^{18} (~ 2^{60})	exabyte	EB	10^{18} B
zettabit	Zbit	10^{21} (~ 2^{70})	zettabyte	ZB	10^{21} B
yottabit	Ybit	10^{24} (~ 2^{80})	yottabyte	YB	10^{24} B

I multipli dell'unità di misura bit sono importanti nelle misure delle trasmissioni dati digitali come la velocità di download e upload in rete mentre i multipli delle unità di misura byte sono importanti nel ridimensionamento dei dispositivi di archiviazione dati.

Un CODICE è un sistema di simboli che permette la rappresentazione dell'informazione ed è definito sulla base di un simbolo che rappresenta l'elemento atomico, di un alfabeto che rappresenta l'insieme dei simboli possibili, di una stringa che rappresenta una sequenza possibile di simboli ed infine un linguaggio che rappresenta una sequenza possibile di stringhe.

Indichiamo con **cardinalità del codice** il numero di elementi dell'alfabeto.

Per passare da un codice all'altro è necessario determinare delle regole di corrispondenza dette codifiche. La codifica mette in corrispondenza biunivoca ogni simbolo con una stringa di simboli. La **notazione numerica è posizionale in quanto ogni simbolo assume un valore più o meno significativo a seconda della posizione nella stringa di numeri.**

Notazione decimale base 10: occorrono 10 simboli (u, da, h, e k)

Il codice binario di base 2: un calcolatore rappresenta l'informazione digitale attraverso la codifica binaria che rappresenta i numeri in modo **posizionale**. La codifica binaria rappresenta numeri mediante una sequenza di 0 e 1.

Ogni elemento di una sequenza viene detto bi.

Una sequenza di 8 bit viene detto byte.

1100 (2, base binario) = **2 simboli (0;1)**

1(due alla terza)1(due alla seconda)0(due alla prima)0 (due alla zero).

Le codifiche numeriche sono posizionali e l'essere umano usa la notazione decimale in base 10 mentre i calcolatori usano delle codifiche digitali in base 2 o multipli di 2 con una codifica binaria o esadecimale.

Un testo è una struttura linguistica molto complessa, contenente informazioni articolate su più livelli. La rappresentazione digitale di un testo detta anche codifica del testo si articola in due livelli:

- Codifica dei caratteri (ad ogni carattere viene associato un codice numerico binario)
- Codifica di alto livello, tramite tag e linguaggi di marcatura xml

La differenza dei caratteri presenti nelle varie lingue pone problemi di compatibilità per la gestione del testo digitale. Per sistemare le cose esistono organismi di standardizzazione per impostare sistemi di codifica univoci.

Il codice ASCII venne ideato nel 1961 ed approvato nel 1968 dall'ANSI, accettato poi come **ISO/IEC 646**

Usa codici binari a 7 bit (128 caratteri codificabili)

I primi 32 codici sono associati a caratteri di controllo non stampabili ed il carattere "spazio" è il primo stampabile. I segni stampabili rimanenti sono associati a: **segni di interpunzione, numeri, caratteri alfabetici minuscoli e maiuscoli.**

CODICE ISO LATIN-1 (ISO 8859-1)

È la codifica dei caratteri standard per il trattamento informatico del testo a 8 bit=1 byte e i caratteri dell'alfabeto latino sono codificati nell'ISO 8859-1, per differenziarlo dagli altri 14 set di caratteri comunemente usati a computer.

Usa codici a 8 bit con 256 caratteri codificabili

I primi 128 caratteri corrispondono all'ASCII

I codici da 128 a 159 sono caratteri di controllo grafico

I codici rimanenti da 160 a 255 sono caratteri alfabetici europei.

CODICE UNICODE

È un sistema di codifica che assegna un numero univoco ad ogni carattere usato per la scrittura di testi. Nota come codifica a 16 bit con più di 65 mila caratteri modificabili, allineato alla standard iso/iec 10646 con una codifica fino a 21 bit.

I primi 256 caratteri corrispondono all'ISO LATIN-1

I codici rimanenti sono associanti agli altri alfabeti conosciuti ed usati dall'uomo.

CODICE UTF-8

Codifica dei caratteri in sequenze di lunghezza variabile di byte creata nel 1992 creata da ROB PIKE. Usa gruppi di byte per rappresentare i caratteri Unicode ed è usata per il trasferimento dei caratteri sul web. Permette di rappresentare tutti i caratteri Unicode a differenza di codifiche più vecchie

I sistemi di codifica dei caratteri non prendono in considerazione l'aspetto grafico dei caratteri. Per questo motivo sono stati inventati i **font** che servono a caratterizzare esteticamente le proprietà visive di un alfabeto. Solitamente i font sono memorizzati in file o possono risiedere su memorie fisse come delle stampanti.

Font fissi: in formato raster (ad ogni carattere corrisponde un'immagine bitmap)

Font scalabili: in formato vettoriale (per ogni carattere le informazioni si riferiscono alla geometria ed ai parametri del disegno).

I documenti di testo possono essere salvati, modificati e trasformati in file di differente "formato" ed alcuni di essi fanno riferimento a software specifici proprietari mentre altri sono generici o standardizzati:

XML E HTML

La struttura di un testo complesso viene descritta attraverso una codifica di alto livello che è associata ad un linguaggio di marcatura detto "markup" in cui porzioni di testo vengono marcate da etichette che caratterizzano la porzione di testo stesso.

I testi non formattati sono i più semplici da codificare e sono costituiti dalla sequenza dei codici del charset scelto.

ESTENSIONE: .TXT

Le possibilità di impaginazione derivano dall'uso dei caratteri di controllo non stampabili. Vengono riconosciuti da tutti gli editor e sistemi operativi.

Un testo complesso che contenga una struttura in sezioni può essere reso visivamente secondo uno stile grafico.

I formati di codifica possono essere **proprietary (con codici stabiliti per un applicativo specifico e con scarsa possibilità di interscambio con altri software); standardizzati (con parole di controllo definite tramite codifica testuale, di maggiori dimensione nei file e con ampia possibilità di interscambio)**

I formati di videoscrittura più usati:

RTF (linguaggio di descrizione interpiattaforma con un numero limitato di parametri di descrizione)

POSTSCRIPT (linguaggio standard di descrizione della pagina in grado di codificare anche informazioni grafiche)

PDF (formato sviluppato nel 1993 per ampliare la qualità e la portabilità del linguaggio postscript con funzioni di ricerca e navigazione)

DOC/DOCX (linguaggio proprietario in grado di codificare molte informazioni ma presenta scarsa compatibilità con altri software)

LATEX (diffuso in ambito scientifico e tecnico per espressioni matematiche)

HTML (linguaggio di descrizione standard, multiplatforma in grado di codificare anche collegamenti ipertestuali)

IMMAGINE DIGITALE

VETTORIALE: generata al computer mediante formule matematiche e descritta mediante un insieme di primitive geometriche

RASTER (bitmap): l'immagine è formata da una griglia di pixel ovvero l'elemento più piccolo che costituisce l'immagine.

Dopo la lavorazione le immagini vettoriali vengono esportate o convertite in formato raster per poterle pubblicare o utilizzare in programmi terzi.

I monitor dei dispositivi che utilizziamo sono strutturati a griglia di pixel proprio come le immagini raster.

DPI (DOTS PER INCH)

- **Numero di punti per pollice;**
- **Misura la densità di punti all'interno di un determinato spazio;**
- **Si fa riferimento ai punti di stampa per digitalizzare un'immagine.**

PPI (PIXEL PER INCH)

- **Numero di pixel per pollice**
- **È il rapporto tra i pixel di un'immagine e i pollici della diagonale del monitor**
- **Si ragiona in ppi quando lavoriamo in un ambiente digitale per poi passare a DPI nel momento della stampa**

PROFONDITA' COLORE è il numero di bit usati per rappresentare il colore di un singolo pixel in un'immagine. (bbp: bit per pixel)

Maggiore è il numero di bit, maggiore sarà la gamma di tonalità che si possono rappresentare.

Modello colore: è il modello matematico che permette la rappresentazione dei colori nel mondo digitale = **RGB E CYMK**

RGB:

- **Modello colore di tipo additivo**

- **Tricomia**
- **Colori ricavati dalla somma delle luminosità dei tre canali**
- **Progettato per la visione su monitor**

CYMK:

- **Modello colore di tipo sottattivo**
- **Clan, magenta, yellow and keyblack**
- **Colori ricavati dalla sottrazione dei 4 colori primari;**
- **Progettato per la stampa fotografica**

Spazio colore: all'interno di un modello colore troviamo diversi tipi di spazio colore, ossia la gamma limitata di colori che sono rappresentabili a seconda della funzione utilizzata.

Profilo colore ICC: ogni dispositivo che registra o produce immagini può avere un profilo ICC associato, che descrive il modo in cui i colori sono rappresentati in un determinato spazio dei colori il cui scopo pratico è quello di consentire l'applicazione di un processo di gestione del colore che preservi i colori nel passaggio da uno spazio colore sorgente ad uno spazio colore destinazione.

Compressione: per ridurre le dimensioni dei file le immagini vengono compresse:

- **Compressione Lossless (compressione senza perdita)**
- **Compressione lossy (compressione con poca perdita)**
- **Non compresso**

FORMATI:

- Bmp (non compresso)
- Raw (nprn compresso)
- Png (possibilità di salvarlo sia compresso che non)
- Tga (senza perdita)
- Tiff (diversi tipi di compressione)
- Gif (senza perdita fino a 256 colori, con oltre)
- Jpg (con perdita)

VIDEO DIGITALE E RISOLUZIONE

Il pixel consiste nell'elemento più piccolo che costituisce l'immagine.

Le differenti risoluzioni video sono:

- SD (standard definition 720x576) di tv e dvd
- HD con 1280x720 per i canali hdtv
- FULL HD con 1920x 1080 per il Bluray
- UHD 3840x2160
- 4K 4096x2160
- 8K

La **RISOLUZIONE CROMATICA** definisce quanti e quali colori sono usati per rappresentare l'immagine

Il **SUBSAMPLING** consiste nel sottocampionamento in cui le immagini vengono codificate riservando maggiore risoluzione alla **LUMINANZA** rispetto all'**CROMINANZA**.

AUDIO DIGITALE

Il CAMPIONAMENTO avviene prelevando dei campioni del segnale originario a intervalli di tempo regolari nella trasformazione del segnale da analogico a digitale.

La FREQUENZA DI CAMPIONAMENTO è la misura del numero di volte al secondo in cui il segnale viene campionato e memorizzato in forma digitale. (hertz)

Per convertire il segnale analogico in forma binaria, ovvero digitale, è necessaria l'operazione di quantizzazione che approssima il valore reale con un valore quantizzato.

Questa approssimazione introduce un errore chiamato ERRORE DI QUANTIZZAZIONE.

Per ridurre le dimensioni, il file audio viene compresso così come accade per le immagini digitali con gli stessi metodi di compressione LOSSLESS, LOSSY E NON COMPRESSO.

CODEC:

SENZA COMPRESSIONE:

- AIFF
- WAV

COMPRESSIONE LOSSLESS

- APPLE LOSSLESS ENCODING
- FLAC
- IFF

COMPRESSIONE LOSSY

- AAC
- DOLBY DIGITAL
- MP3
- REALAUDIO
- WINDOWS MEDIA AUDIO

**CAMPIONAMENTO(hertz) + DURATA + COMPRESSIONE(bitrate) = PESO/
QUALITA'**

L' MP3 è l'algoritmo di compressione di tipo LOSSY, che elimina dai file audio determinate informazioni che non sono necessarie (come le frequenze non udibili dall'orecchio umano).

Il MIDI è uno standard creato per lo scambio di dati fra strumenti musicali elettronici a prescindere dalla marca e dal modello.

MIDI VS WAVEFORM

- In un file midi sono memorizzati solo dei segnali come una nota premuta;
- I file MIDI sono molto più piccoli dei file WAVEFORM anche da compressi;
- Dal midi si può ricostruire esattamente uno spartito
- I file midi possono essere modificati nota per nota
- E suona in modo diverso a seconda dell'hardware e del software di riproduzione

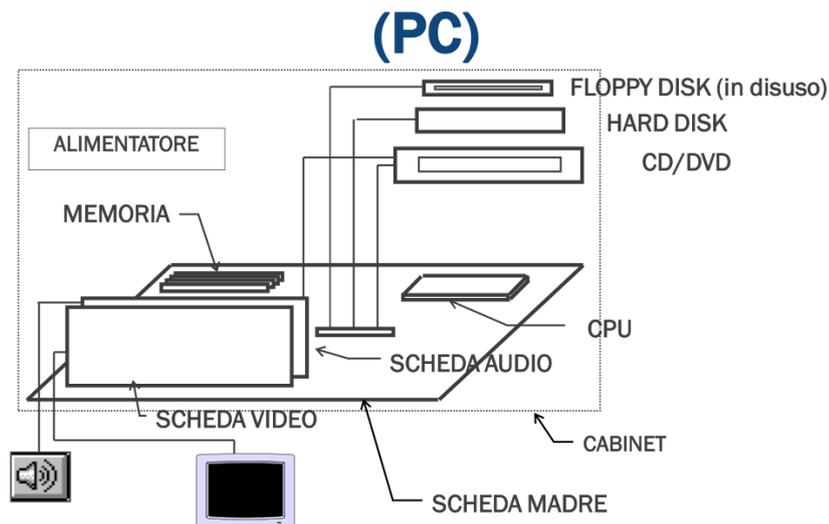
IL CALCOLATORE

Un calcolatore è una macchina «esecutrice» in grado di elaborare e memorizzare automaticamente informazioni.

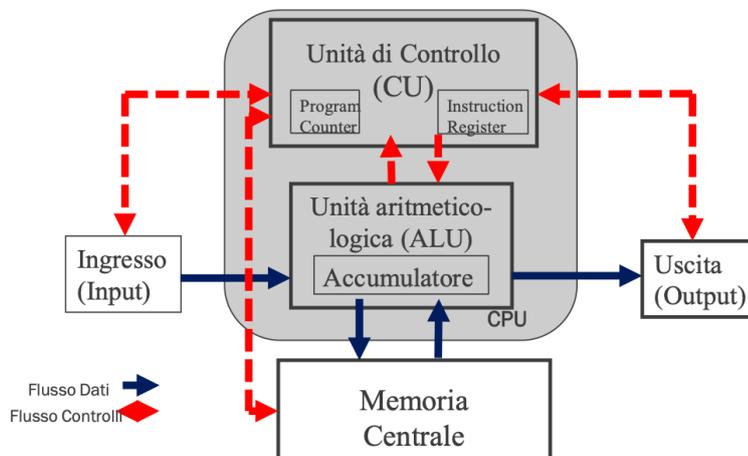
- Un calcolatore può elaborare informazioni rappresentate in un formato detto digitale.
- La caratteristica fondamentale di un calcolatore è quella di essere una macchina programmabile, basata su un «hardware» (componenti elettronici) su cui «girano» (vengono elaborati) «programmi» (software).

IL COMPORTAMENTO DI UN CALCOLATORE segue queste tre fasi:

- Legge un insieme di informazioni come dati input
- Elabora queste informazioni attraverso un insieme di istruzioni con un programma
- E restituisce le informazioni risultato dell'elaborazione in dati di uscita



Macchina di Von Neumann (1945)



La CPU dal modello di VON NEUMANN

È il cervello del calcolatore ed esegue i programmi contenuti nel calcolatore comandando le altre parti del calcolatore.

È composta da due parti: la CU che esegue le istruzioni, comanda le altre parti della CPU e controlla il flusso delle operazioni tra CPU e memoria;

La ALU esegue le operazioni aritmetiche, le operazioni logiche e i confronti tra dati.

Il funzionamento della CPU è dato dal seguente CICLO MACCHINA:

FETCH - DECODE - EXECUTE CYCLE

Nella fase di caricamento (fetch) la Central unit preleva istruzioni dalla memoria

Nella fase di decodifica interpreta le istruzioni e trasferisce dalla memoria i dati necessari alla sua esecuzione.

Nella fase di esecuzione comanda le parti ed esegue i calcoli associati all'istruzione decodifica

Nella fase di memorizzazione i risultati vengono memorizzati nella memoria centrale o in registri della CPU.

Il set di istruzioni definisce il linguaggio comprensibile dalla CPU e questo linguaggio viene chiamato LINGUAGGIO MACCHINA. I programmi eseguibili dalla CPU sono sequenza di istruzioni in linguaggio macchina e sono molto semplici: istruzioni di lettura e scrittura nella memoria centrale, istruzioni aritmetiche ed istruzioni di salto ad una certa istruzione del programma.

Le CPU si distinguono in base al tipo di set di istruzioni – CISC (Complex Instruction Set Computer)

– RISC (Reduced Instruction Set Computer)

- La prestazioni di una CPU sono misurate dal numero di istruzioni eseguite al secondo.
- L'unità di misura adottata è il MIPS (= Million Instruction per second) o multipli.

La CPU è costituita da un microprocessore le cui capacità possono essere stimate da

– Numero di transistor

– Frequenza di clock (misurato in hertz – Hz o cicli/sec)

- Più grande è il numero di transistor del processore più complesse sono le operazioni eseguibili a parità di tempo.

- Il clock è il segnale che cadenza l'esecuzione delle singole istruzioni. Quindi più alta è la frequenza di clock più operazioni sono eseguite a parità di tempo.

MEMORIA CENTRALE

- L'informazione è memorizzata per mezzo di singoli dispositivi elettronico che possono trovarsi ognuno in due possibili stati:

– Un interruttore che può essere aperto oppure chiuso.

– Un condensatore che può essere carico oppure scarico.

- La memoria è organizzata in una sequenza logica di locazioni individuate da un indirizzo (address) e contenente una parola (word) di informazione.

– Architetture a 32 bit = la parola (word) è di 4 byte

– Architetture a 64 bit = la parola (word) è di 8 byte

CAPACITA' DELLA MEMORIA

K (kilo)	$2^{10} = 1024 \cong 10^3 = 1000$
M (mega)	$2^{20} \cong 1.000.000$
G (giga)	$2^{30} \cong 1.000.000.000$
T (tera)	$2^{40} \cong 1.000.000.000.000$

- La capacità di una memoria si misura in byte, ma per convenzione si usano le equivalenze approssimate tra le potenze decimali e binarie

- Il tempo di accesso di una locazione di memoria è dell'ordine delle decine di nanosecondi.

- La memoria centrale (NON è destinata a contenere grandi moli di dati che devono essere conservati in modo permanente) di un calcolatore contiene tre tipi di memoria elettronica

– RAM

- Riscrivibile

- Casuale

- Volatile

- - ROM

- Non riscrivibile

- Casuale

- Non volatile

– CACHE (il tipo di memoria di transito veloce tra CPU e memoria principale, caratterizzata in livelli a seconda del rapporto dimensione/velocità/costo.)

LE UNITA' PERIFERICHE

- Le unità periferiche permettono al calcolatore

– La comunicazione con il mondo esterno.

– Il mantenimento di grandi quantità di informazioni non gestibili dalla memoria centrale.

- Le unità periferiche si dividono in tre classi

– Ingresso (Input) e permettono al calcolatore di ricevere informazioni dal mondo esterno e sono basate su sensori o parti elettroniche (mouse, tastiera, touchpad, scanner, microfono)

– Uscita (Output) permettono al calcolatore di inviare informazioni all'esterno e sono basate su parte elettronica per mantenere l'informazione, per codificarla o per segnalare lo stato dell'unità oppure sono basate su attuatori che trasformano il segnale elettrico digitale in posizione, temperatura e luminosità. (monitor, stampante, casse acustiche)

– Input/Output possono essere utilizzate indipendentemente come ingresso e uscita (hard-disk, usb, cdrom).

Le unità (drive) a dischi magnetici (HDD) sono unità di I/O che usano la polarizzazione di un materiale ferromagnetico (ossido di ferro) depositato su un supporto inerte.

– Riscrivibili – Non volatili

- Le unità (drive) disco a stato solido (SSD) sono memoria di massa basate su semiconduttori (memorie flash)

– Riscrivibili – Non volatili