

# **Il modello di esecutore**

## Capitolo II

# Trattamento delle Informazioni

- Informatica = studio sistematico dei *processi* che servono al trattamento delle informazioni o più in generale della definizione della soluzione di problemi assegnati
  - analisi dettagliata di ciò che serve al trattamento dell'informazione
  - progetto di una soluzione applicabile alla generazione di informazioni prodotte da altre informazioni
  - correttezza e efficienza della soluzione pensata

# Informatica e studio di Algoritmi

- “algoritmo”
  - la sequenza precisa di operazioni il cui svolgimento sequenziale è necessario per la soluzione di un problema assegnato
- Informatica → studio sistematico degli algoritmi
  - Il calcolatore è tra tutti gli esecutori di algoritmi (compreso l'uomo) quello che si mostra più potente degli altri e con una potenza tale da permettere di gestire quantità di informazioni altrimenti non trattabili

# Algoritmo ed Esecutore

- Algoritmo
  - un testo che prescrive un insieme di operazioni o azioni eseguendo le quali è possibile risolvere il problema assegnato
  - Se si indica con istruzione la prescrizione di una singola operazione, allora l'algoritmo è un insieme di istruzioni da svolgere secondo un ordine prefissato
- Esecutore
  - l'uomo o la macchina in grado di risolvere il problema eseguendo l'algoritmo
  - Se un algoritmo è un insieme di istruzioni da eseguire secondo un ordine prefissato, allora l'esecutore non solo deve comprendere le singole istruzioni ma deve essere anche capace di eseguirle

# Cos'è un computer

- Un computer è un apparecchio elettronico
  - strutturalmente non ha niente di diverso da un televisore, uno stereo, un telefono cellulare o una calcolatrice elettronica
- progettato per eseguire autonomamente e velocemente attività diverse
  - Come tutte le macchine, non ha nessuna capacità decisionale o discrezionale, ma si limita a compiere determinate azioni *secondo procedure prestabilite*
- Il computer è una macchina che in maniera automatica esegue operazioni “*elementari*” ad altissima velocità
  - L'altissima velocità di elaborazione (ad es. milioni di istruzioni per secondo) fa sì che operazioni complicate, espresse mediante un gran numero di operazioni semplici, siano eseguite in tempi ragionevoli per l'ambiente esterno

# Algoritmo e Programma

- Informalmente, un *algoritmo* è la descrizione di un lavoro da far svolgere ad un *esecutore*
  - Se si vuole usare un computer bisogna non solo progettare preliminarmente un algoritmo, ma anche provvedere a comunicarglielo in modo che gli risulti comprensibile
- La descrizione di un algoritmo in un linguaggio comprensibile ad un calcolatore è detto Programma
- Il linguaggio nel quale viene scritto il programma è detto linguaggio di programmazione

# Programma

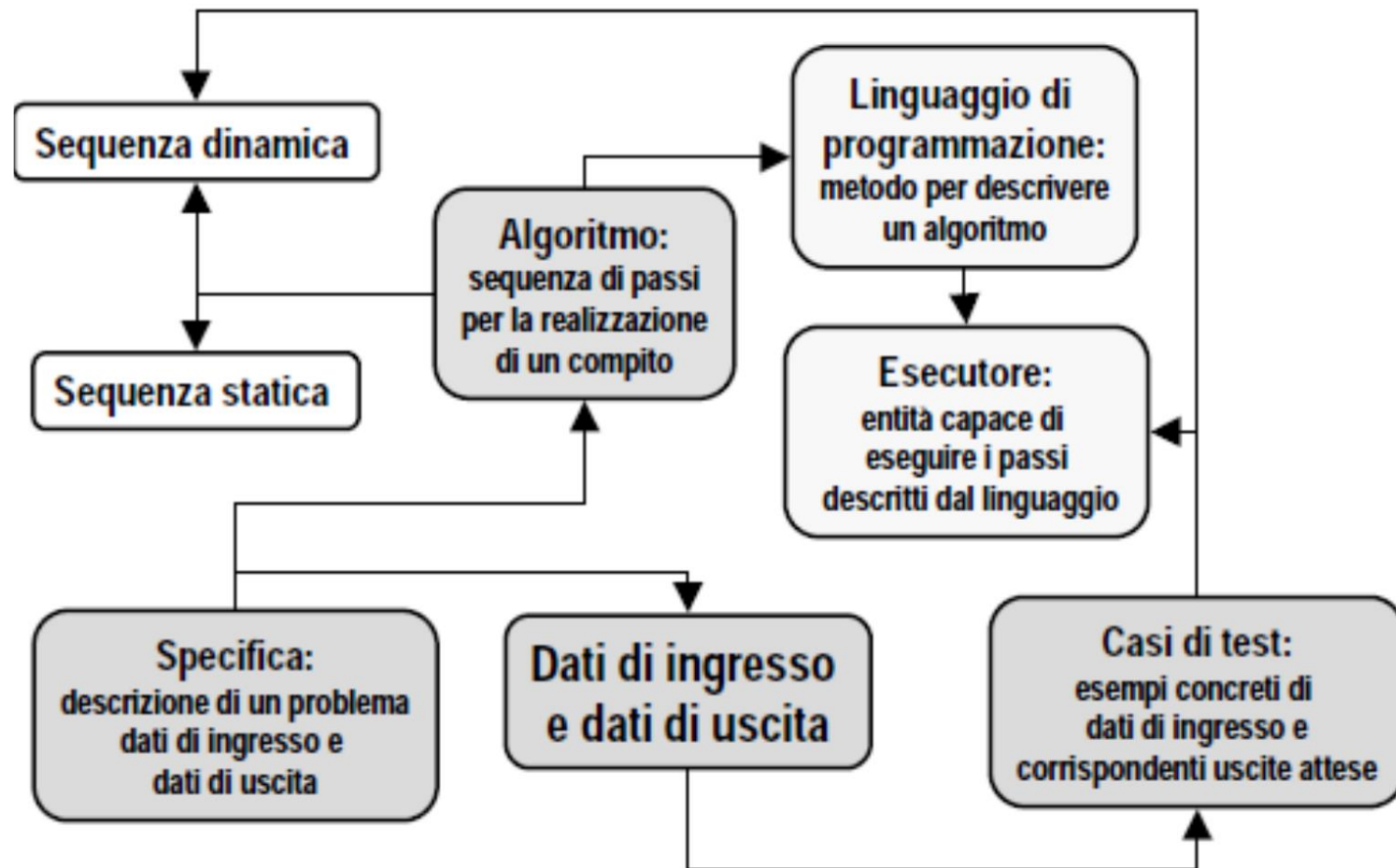
- La specifica è la descrizione in linguaggio naturale del problema da risolvere
- Informazioni di ingresso
- Informazioni di uscita
- Esempi:
  - Somma di due numeri naturali di una cifra
    - Input: i due numeri; output: la somma e il riporto
  - Somma di due numeri naturali in colonna
    - Input: i due numeri naturali; output: la somma

# Processo e Processore

- Si definisce *processo* il lavoro svolto eseguendo l'algoritmo, e *processore* il suo esecutore
  - Il processo non è altro che l'elenco delle azioni effettivamente svolte come si susseguono nel tempo
  - Sequenza statica e sequenza dinamica

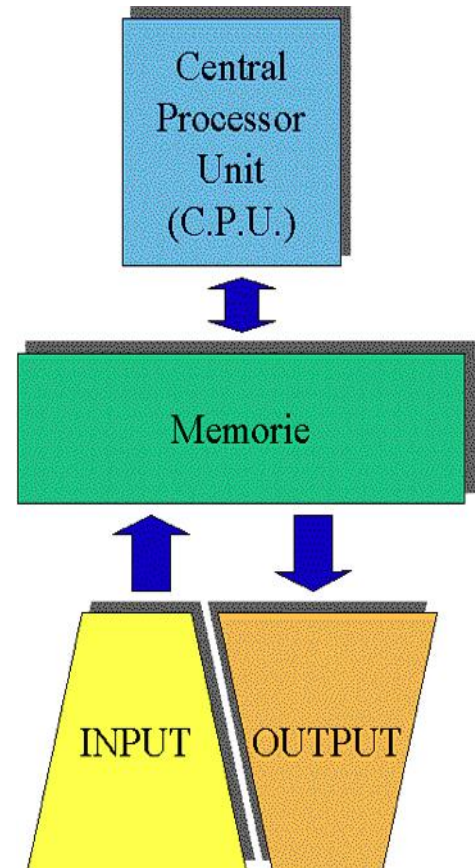


# Riepilogo



# Il modello di Von Neumann

- E' uno schema di principio rappresentativo dei tradizionali computer
- Prende il nome da Von Neumann, il primo ricercatore che lo propose nel 1945
- La *Central Processing Unit (CPU)* coordina l'esecuzione delle operazioni fondamentali
- La *memoria* contiene un programma che descrive le operazioni da eseguire e i dati su cui il programma stesso opera
- I dispositivi di *input* e *output* sono le interfacce della CPU nei confronti del mondo esterno



# Stored Procedure

- Il modello di Von Neumann si basa sul concetto di programma memorizzato
  - la macchina immagazzina nella propria memoria i dati su cui lavorare e le istruzioni per il suo funzionamento
- Da cui ne deriva una grande flessibilità operativa
  - macchine nate per fare calcoli possono essere impiegate nella risoluzione di problemi di natura completamente diversa, come problemi di tipo amministrativo, gestionale e produttivo

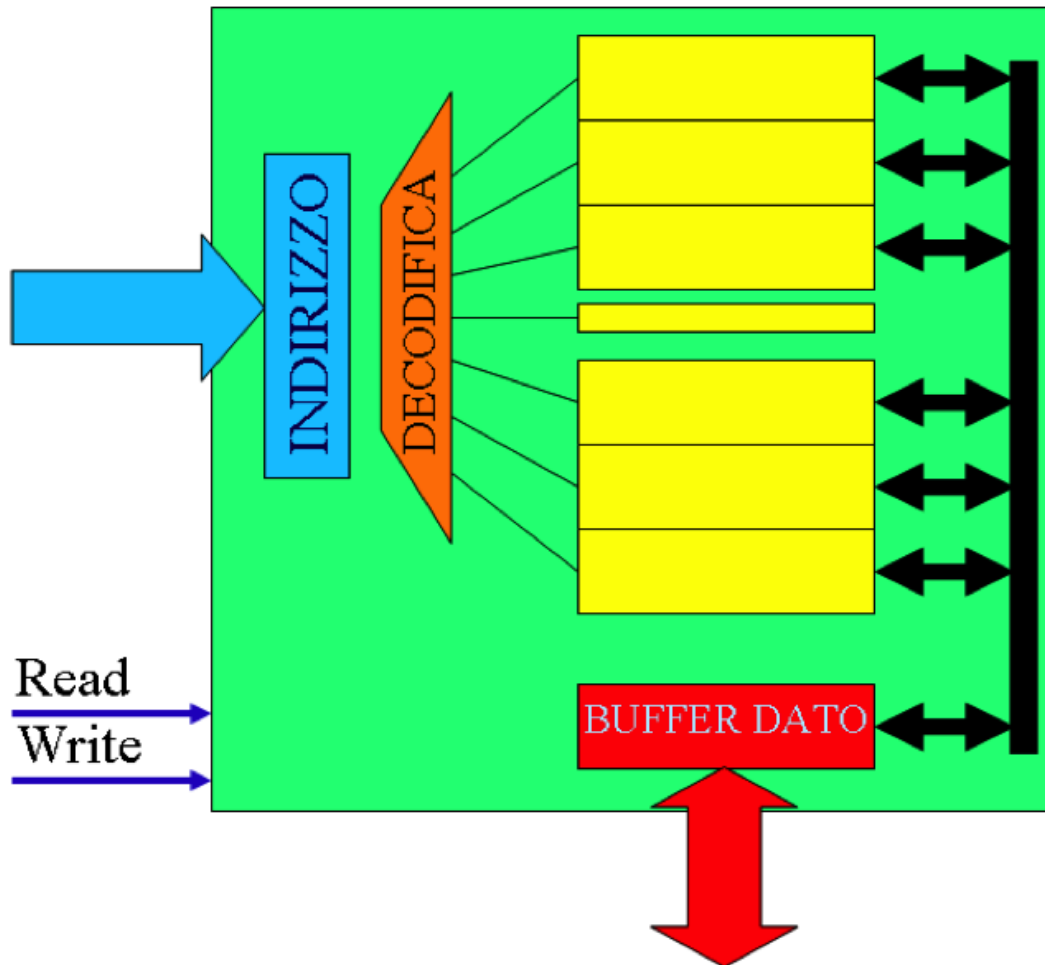
# Funzionamento interno

- Nelle linee generali il funzionamento interno di un qualsiasi computer si può ricondurre al modello di Von Neumann che non è molto dissimile da uno antropomorfo che vede un essere umano
  - sostituire con il suo cervello la CPU
  - fogli di carta alla memoria centrale
  - una calcolatrice con le operazioni fondamentali all'*A.L.U* (*Aritmetic Logic Unit*), ovvero il componente della CPU che effettua tutte le operazioni di calcolo logico e aritmetico

# Memorie

- In generale le memorie possono essere viste come un insieme di contenitori fisici, detti anche *registri*, di dimensioni finite e fissate a cui si può far riferimento mediante la posizione occupata nell'insieme detta *indirizzo* di memoria
  - La dimensione di un registro si misura in numero di bit
  - I registri per i singoli bit possono essere fatti in diversi modi
    - nelle memorie di tipo elettronico sono circuiti detti flip-flop che mostrano un valore di tensione o uguale a 5 Volt o a 0 Volt, nelle memorie di tipo magnetico è una sorta di calamita polarizzata o positivamente o negativamente, nelle memorie di tipo ottico è una superficie con o senza un buco in modo da riflettere diversamente il raggio laser che la colpisce
  - In ogni caso il dispositivo di lettura deve essere in grado di associare allo stato del bit il valore 1 (ad esempio tensione a 5 volt, polo positivo, assenza di buco) o il valore 0 (tensione a 0 volt, polo negativo, presenza di buco)
- Memorie con registri di otto bit sono dette a *byte o caratteri*; con più di otto (solitamente 16 o 32) vengono invece dette *a voce*
  - I calcolatori moderni sono dotati di memorie a byte e le memorie a voce sono solo un ricordo del passato

# Schema



# Operazioni sulla memoria

- Load
  - si preleva l'informazione contenuta nel registro senza però distruggerla
- Store
  - si inserisce una informazione nel registro eliminando quella precedente
- Per comprendere il funzionamento di un registro di memoria si può pensare ad una lavagna il cui uso può essere così esemplificato
  - leggere informazioni a patto che vi siano state scritte
  - la lettura non cancella quanto scritto
  - la scrittura di nuove informazioni obbliga a cancellare quelle precedenti che pertanto vengono perse
- La memoria è un sistema che assolve al compito di conservare il dato, depositandolo in un registro nel caso di operazione di scrittura, e di fornire il dato conservato in un registro in caso contrario di operazione

# Funzionamento di una Memoria

- La CPU indica preventivamente l'indirizzo del registro interessato dall'operazione
- La memoria decodifica tale indirizzo abilitando solo il registro ad esso corrispondente affinché
  - per uno *store* copi il dato del buffer nel registro
  - per un *load* copi il dato del registro nel buffer
    - dove il *buffer* può essere vista come un'area di transito dei dati dalla CPU alla memoria e viceversa
- Le operazioni di load e store richiedono tempi di attuazione che dipendono dalle tecnologie usate per la costruzione delle memorie e dalle modalità di accesso
  - Nel caso di load, il tempo di accesso misura il tempo che trascorre tra la selezione del registro di memoria e la disponibilità del suo contenuto nel registro di buffer
  - Il tempo di accesso nel caso dello store misura invece il tempo necessario alla selezione del registro e il deposito del contenuto del registro di buffer in esso

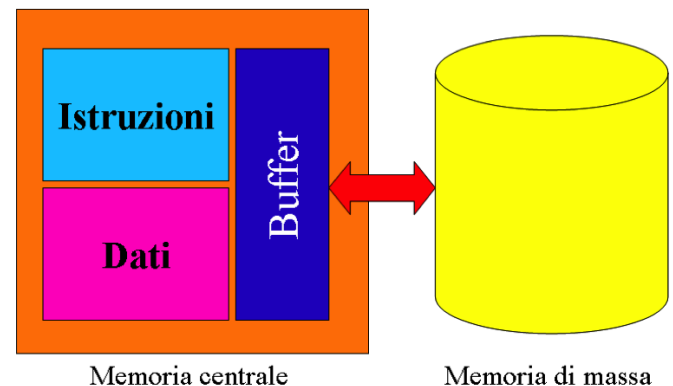
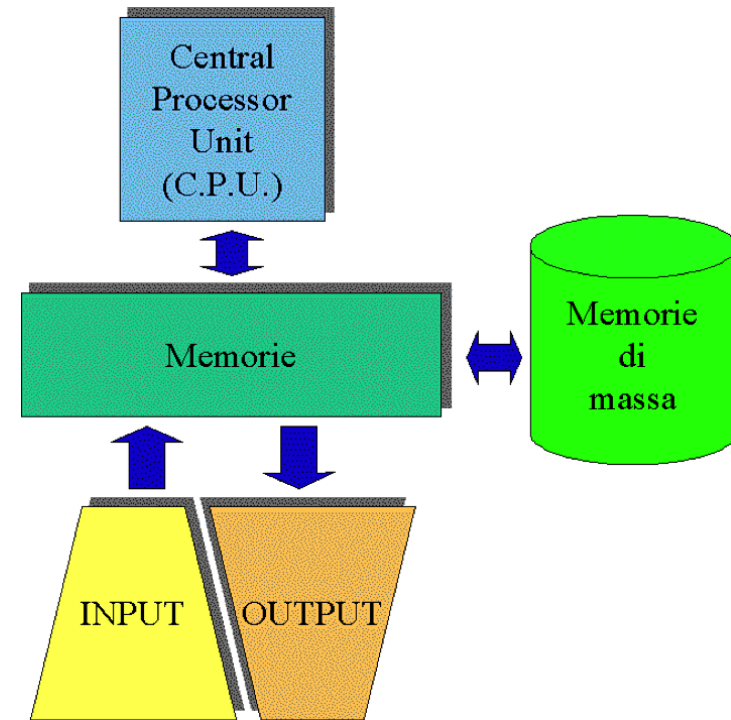


# Diversi tipi di memoria

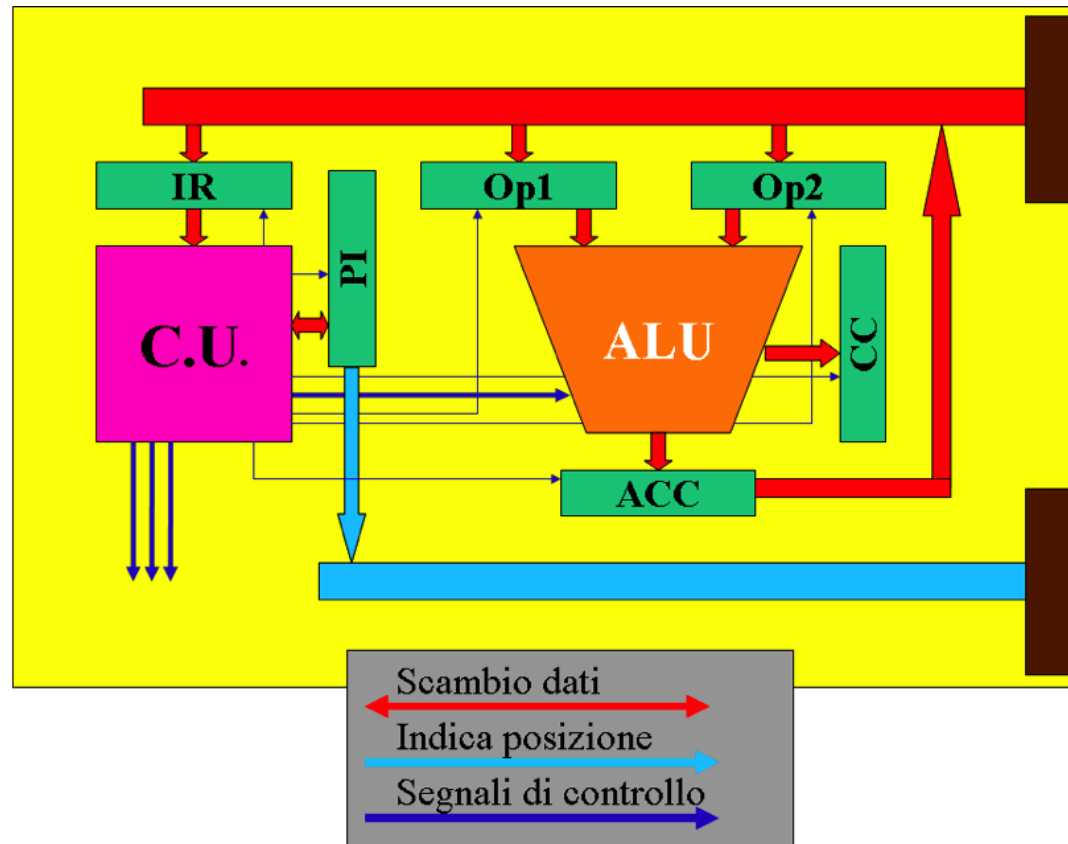
- *Memoria ad accesso casuale*
  - Il tempo di accesso non dipende dalla posizione - *R.A.M.* (Random Access Memory)
- *Memoria ad accesso sequenziale*
  - Il tempo di accesso dipende dalla posizione, come avviene nei nastri magnetici
- Alcune memorie vengono realizzate in modo che sia possibile una sola scrittura di informazioni. Tali memorie vengono dette a *sola lettura o ROM* (da *Read Only Memory*)
  - L'uso di queste memorie è necessario quando si desidera che alcune istruzioni o dati non siano mai alterati o persi
- *Memorie volatili*
  - Perdono le informazioni in esse registrate quando il sistema viene spento
  - Memorie elettroniche
- *Memorie permanenti*
  - Memorie di tipo magnetico, ottico e tutti i tipi di ROM

# Memorie di Massa

- Le memorie di massa sono memorie ausiliarie caratterizzate da una elevata capacità
  - Le informazioni contenute nella memoria di massa devono essere dapprima trasferite nella memoria centrale e successivamente elaborate
  - Le informazioni prodotte dalla CPU, viceversa, devono essere depositate in memoria centrale per poi essere conservate nelle memorie di massa
- Le memorie di massa hanno tempi di accesso maggiori dovuti alle tecnologie impiegate per realizzarle
- Per ovviare alla differenza di velocità tra i due dispositivi si impiegano tecniche che prevedono che la memoria centrale contenga anche *aree di accumulo dei dati in transito verso tutti i dispositivi esterni* (buffer)



# La C. P. U.

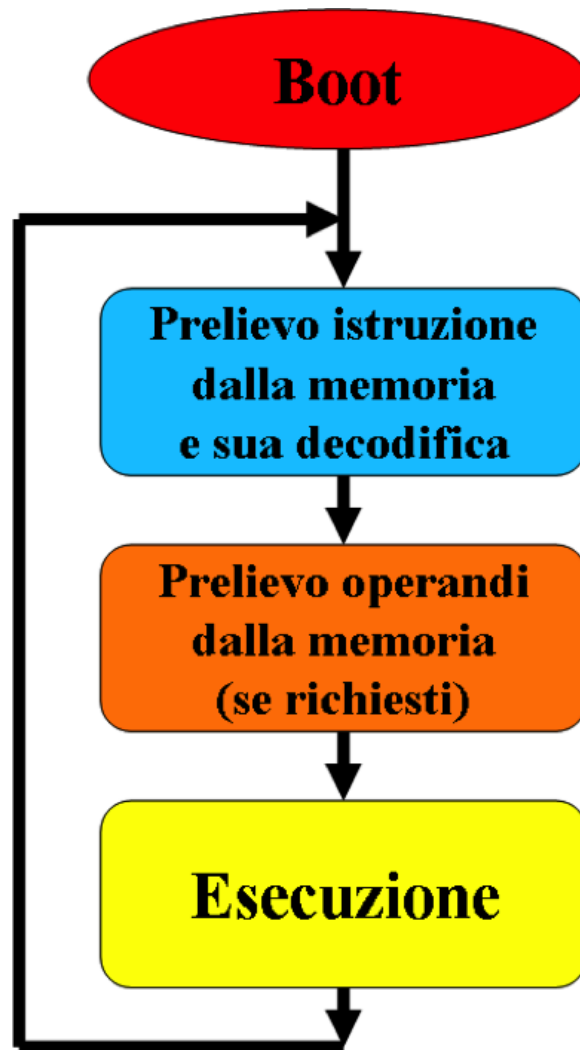


- ALU: unità logica aritmetica
- CU: Unità di Controllo
- Registri Interni

# Control Unit

- È preposto all'interpretazione delle singole istruzioni e all'attivazione di tutti i meccanismi necessari per eseguire le singole istruzioni
- In particolare la CU ha il compito di
  - Prelevare ogni istruzione dalla memoria centrale e la decodifica
  - Prelevare i dati dalla memoria se servono all'istruzione
  - Eseguire l'istruzione
    - Per esempio: se l'istruzione prelevata è di tipo aritmetico e richiede due operandi, la CU predispone dapprima il prelievo dalla memoria di tali operandi, attiva poi l'ALU perchè esegua l'operazione desiderata, ed infine deposita il risultato di nuovo in memoria
- Al termine dell'esecuzione di una istruzione la CU procede al prelievo dalla memoria della successiva istruzione secondo un ordine rigidamente sequenziale
  - L'esecuzione di una istruzione può avere inizio solo se la precedente è stata portata a termine

# Il Ciclo del Processore

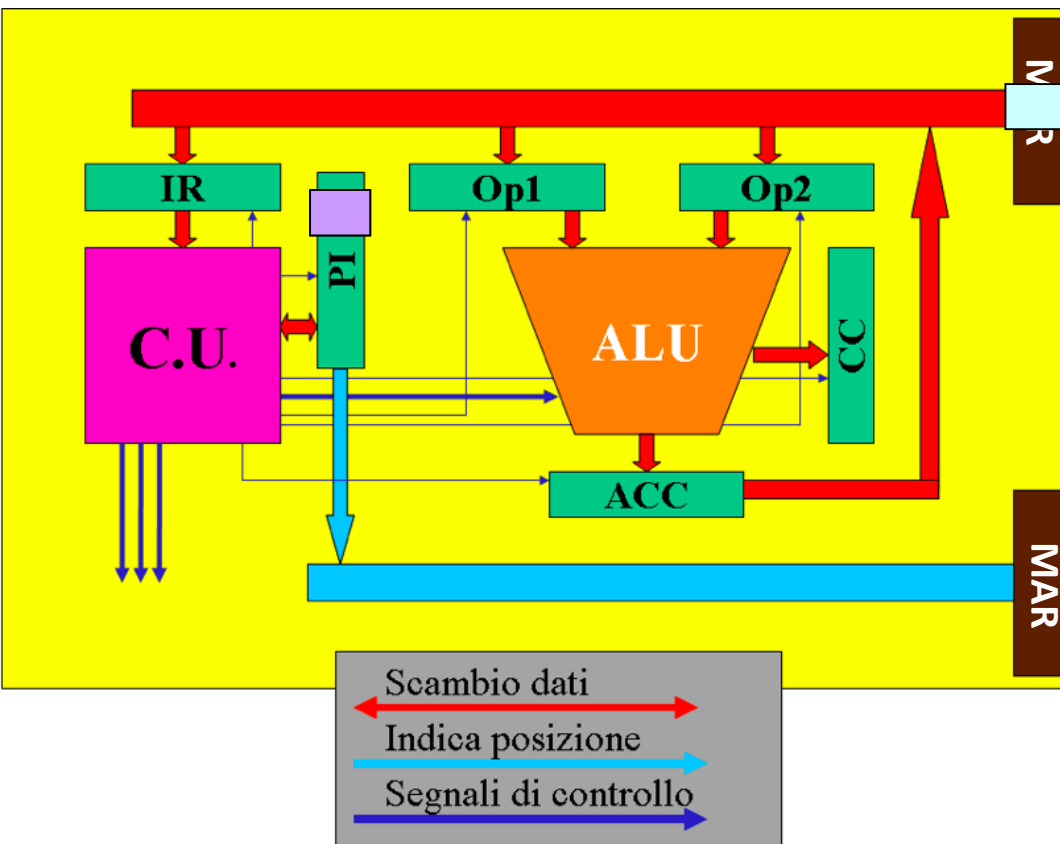


- Le tre fasi del ciclo vengono anche dette
  - *fetch*
  - *operand assembly*
  - *execute*

# Fetch

## Fase di fetch

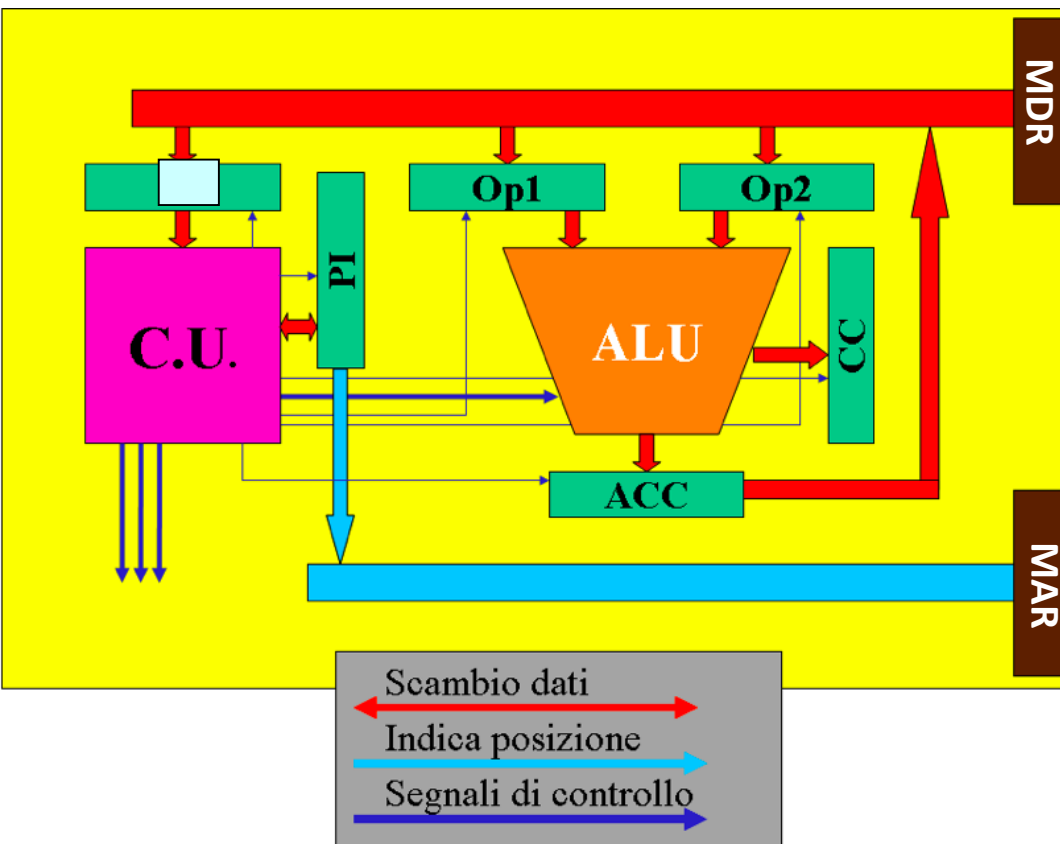
1. Per avviare un programma, residente nella memoria principale, bisogna inserire l'indirizzo della prima istruzione nel PI.
2. La CU ricopia il contenuto del PI nel MAR e viene effettuato un accesso in lettura a quell'indirizzo.
3. Il dato prelevato dall'indirizzo, cui si è acceduto, viene conservato in MDR.
4. Il contenuto di MDR viene ricopiato in IR. Nel frattempo PI si è, in maniera automatica, aggiornato alla locazione successiva.



# Decode

## Fase di decode

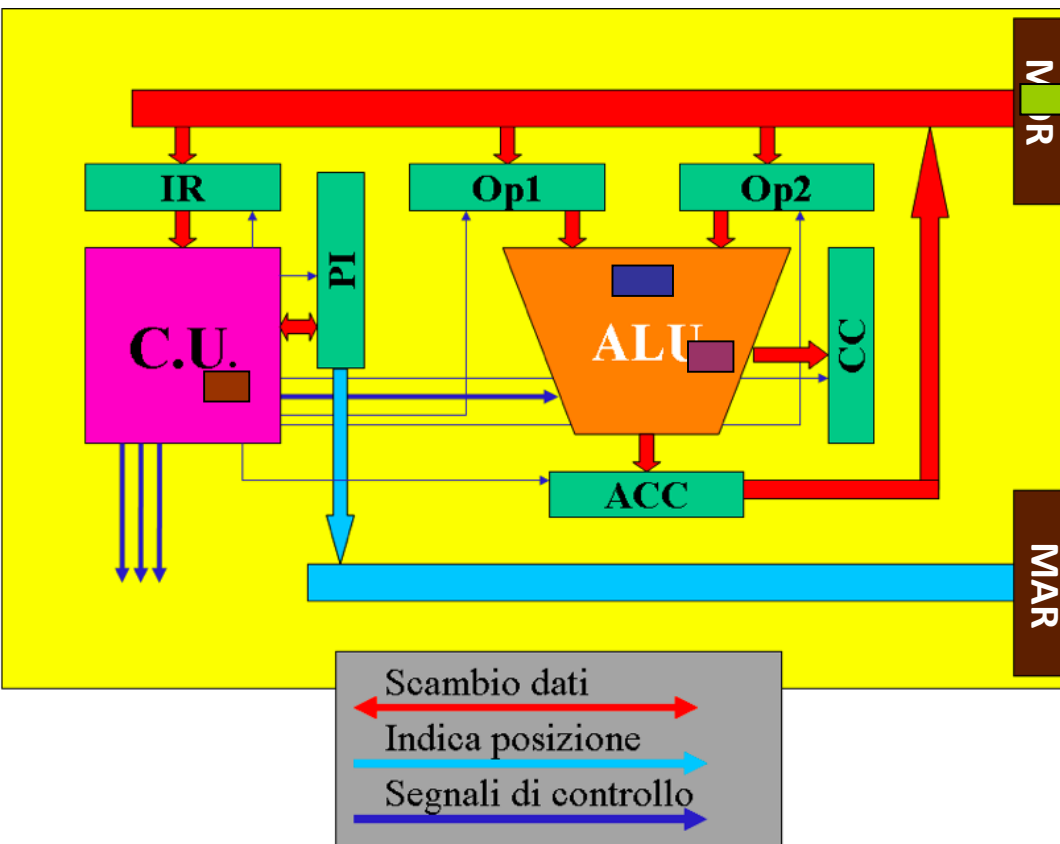
L'istruzione viene interpretata, prelevando eventualmente dalla memoria altri dati necessari se l'istruzione completa non è contenuta nel dato prelevato in precedenza. In ogni caso, alla fine, PI punterà all'istruzione da eseguire successivamente..



# Execute

## Fase di execute

Se, per esempio, l'istruzione da eseguire è una somma, i registri verranno caricati con i numeri da sommare e la CU invia alla ALU il comando per l'esecuzione della somma. La ALU esegue l'operazione: il risultato prodotto sarà conservato in uno dei registri e verrà aggiornata il CC.





# Aritmetic Logic Unit

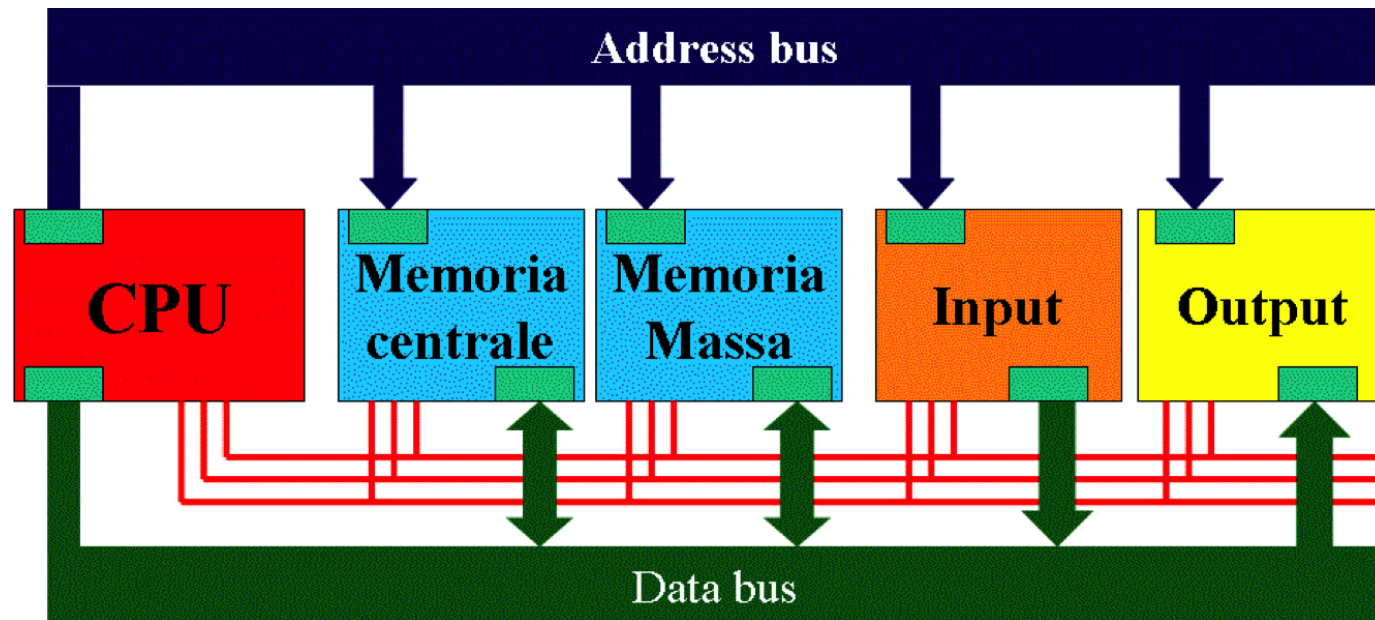
- Esegue operazioni
  - aritmetiche, di confronto o bitwise
  - sui dati della memoria centrale o dei registri interni
- L'esito dei suoi calcoli viene segnalato da appositi bit in un registro chiamato *Condition Code*
- A seconda dei processori l'ALU può essere molto complessa
- Nei sistemi attuali l'ALU viene affiancata da processori dedicati alle operazioni sui numeri in virgola mobile detti *processori matematici*
- Durante le sue elaborazioni la CU può depositare informazioni nei suoi registri interni in quanto sono più facilmente individuabili e hanno tempi di accesso inferiori a quelli dei registri della memoria centrale

# Registri Interni

- Il numero e tipo di tali registri varia a seconda dell'architettura della CPU
- In molte CPU troviamo
  - *Instruction Register (IR)*
    - Contiene l'istruzione prelevata dalla memoria e che l'unità di controllo sta eseguendo
  - *Prossima Istruzione (PI)*
    - Ricorda alla CU la posizione in memoria della successiva istruzione da eseguire
    - Nei casi in cui ogni registro di memoria contenga un'intera istruzione, e l'insieme delle istruzioni del programma sia disposto ad indirizzi consecutivi, la CU incrementa di uno il valore contenuto in PI dopo ogni prelievo di una istruzione dalla memoria
  - *Accumulatore (ACC)*
    - serve come deposito di dati da parte dell'ALU nel senso che contiene prima di un'operazione uno degli operandi, e al termine della stessa operazione il risultato calcolato
      - In questo caso i registri *Op1* e *Op2* diventano interni all'ALU
  - *Condition Code (CC)*
    - indica le condizioni che si verificano durante l'elaborazione, quali risultato nullo, negativo e overflow

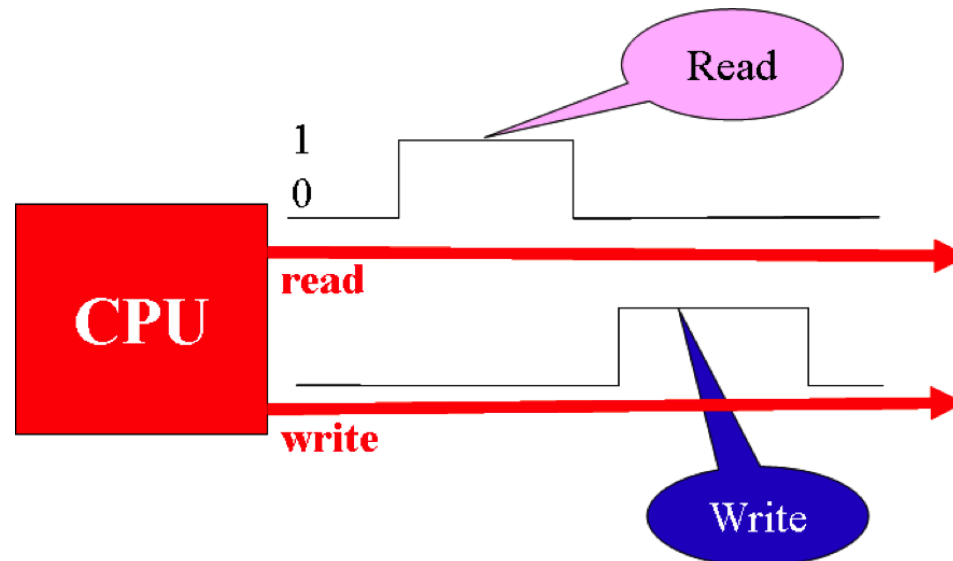
# Il Bus

- E' un canale di comunicazione *condiviso* da più utilizzatori
- E' fisicamente costituito da uno o più fili su cui possono transitare uno o più bit contemporaneamente
- A seconda delle informazioni trasportate si hanno
  - *bus dati (data bus)*
  - *bus indirizzi (address bus)*
  - *bus comandi o di controllo (command o control bus)*



# Lettura e Scrittura

- Il control bus serve alla C.U. per indicare ai dispositivi cosa essi devono fare
  - Tipici segnali del control bus sono quelli di *read* e *write* mediante i quali la CU indica ai dispositivi se devono leggere un dato dal data bus (read) o scriverlo su di esso (write)
  - Il data bus permette ai dati di fluire da CPU a registro di memoria selezionato per operazioni di *store* e viceversa per quelle di *load*



# Address Bus

- L'address bus serve alla CU per comunicare l'indirizzo del dispositivo interessato da una operazione di lettura o scrittura
  - In questa ottica anche i dispositivi di input /output sono identificati da un indirizzo alla stessa stregua dei registri di memoria
- Tutti i componenti del sistema (memoria, input, output, memoria di massa, etc.) devono essere dotati della capacità di riconoscere sull'address bus il proprio indirizzo
  - In altri termini attraverso l'address bus la CU effettua la selezione del dispositivo a cui sono rivolti i comandi e i dati

# Scambio di Informazioni

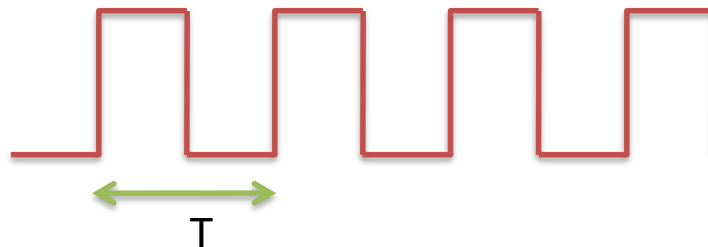
- La CPU è l'unico elemento che fornisce l'indirizzo all'address bus
- Memorie e dispositivi di input ed output devono *ascoltare l'address bus* per attivarsi quando su di esso compare il proprio indirizzo identificativo
  - nel caso della memoria l'attivazione avviene quando viene riconosciuto l'indirizzo corrispondente ad uno dei registri di cui essa è composta
  - il dispositivo attivo deve interpretare i segnali del control bus per eseguire i comandi della CU
  - le memorie prelevano dati dal data bus o immettono dati in esso in funzione del comando impartito dalla CU
- I dispositivi di input possono solo immettere dati sul data bus
- Viceversa i dispositivi di output possono solo prelevare dati dal data bus

# Bus Seriale e Parallelo

- Un bus costituito da un solo filo è chiamato bus *seriale* e su di esso i bit transitano uno dietro l'altro
- Un bus costituito da n fili è chiamato bus *parallelo* perché su di esso transitano n bit alla volta
  - Tipici sono i bus a 8 e 32 fili sui quali si possono trasferire rispettivamente 8 e 32 bit (4 Byte) alla volta
- L'address e il data bus sono paralleli e le loro dimensioni caratterizzano i sistemi di calcolo
  - Il numero di bit dell'address bus indica la *capacità di indirizzamento* della CPU: ossia la sua capacità di gestire la dimensione della memoria centrale e il numero di dispositivi di input ed output
    - Infatti un address bus con n bit consente di selezionare un registro tra  $2^n$
  - La dimensione del data bus condiziona invece la velocità di scambio delle informazioni tra i diversi dispositivi in quanto con m fili solo m bit possono viaggiare contemporaneamente

# Il clock

- Le attività di tutti i dispositivi vengono sincronizzate tra loro mediante un orologio interno che scandisce i ritmi di lavoro
- Il *clock* è un segnale periodico di periodo fisso
  - un'onda quadra caratterizzata da un periodo  $T$  (detto ciclo) e da una frequenza  $f$  ( $f=1/T$ ) misurata in Hertz (Hz)
  - Ad esempio un clock composto da 10 cicli al secondo ha la frequenza  $f = 10$  Hz e il periodo  $T = 100\text{ms}$ 
    - La frequenza dei clock presenti nei moderni sistemi spazia dai MHz (1 MHz corrisponde a un milione di battiti al secondo) ai GHz (1 GHz corrisponde a un miliardo di battiti al secondo)





# Clock e velocità di elaborazione

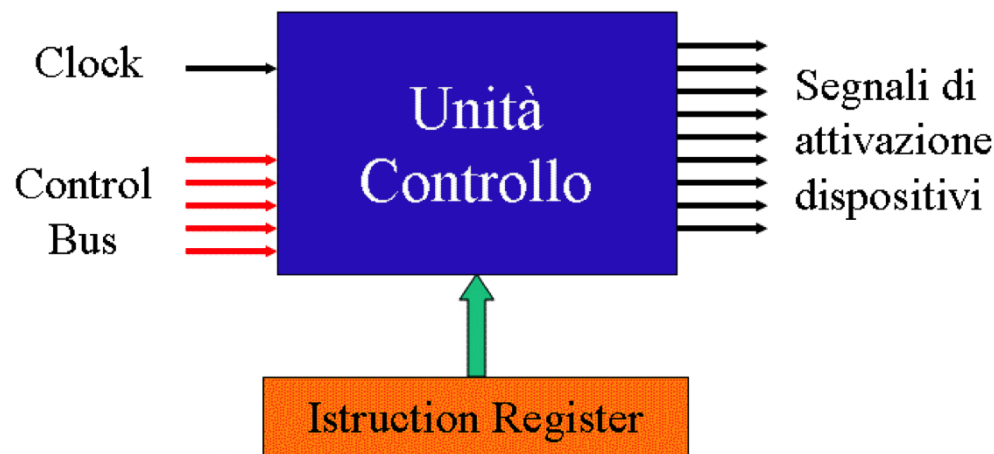
- La velocità di elaborazione di una CPU dipende dalla frequenza del suo clock
  - Più accelerato è il battito del clock maggiore è la velocità di esecuzione
- Alla frequenza del clock è legato il numero di operazioni elementari che vengono eseguite nell'unità di tempo dalla CU
  - Ad esempio, se si assume che ad ogni ciclo di clock corrisponde esattamente l'esecuzione di una sola operazione, allora la frequenza del clock indica il numero di operazioni che vengono eseguite nell'unità di tempo dalla UC
    - Con un clock a 3 GHz si ha che il processore è in grado di eseguire 3 miliardi di operazioni al secondo
    - In realtà tale ipotesi non è sempre vera in quanto l'esecuzione di una operazione può richiedere più cicli di clock sia per la complessità delle operazioni che per la lentezza dei dispositivi collegati alla CPU

# Istruzioni

- Le istruzioni sono operazioni semplici
  - trasferimento dati da un registro ad un altro
    - da memoria a memoria
    - da memoria a registri della CPU o viceversa
    - da memoria a output, da input a memoria
  - operazioni aritmetiche o logiche eseguite dall'ALU
  - controllo di condizioni riportate dal registro CC o deducibili dal confronto di due registri

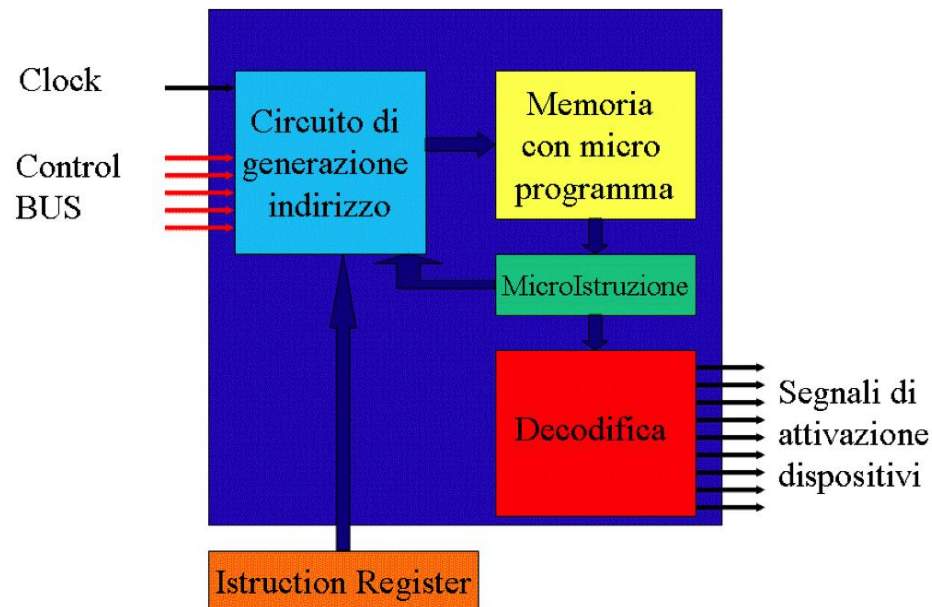
# Esecuzione di una Istruzione

- L'esecuzione di una istruzione da parte della UC consiste *nell'inoltro di una sequenza di abilitazioni* dei dispositivi il cui effetto corrisponde alla operazione richiesta
- Le prime UC erano realizzate con circuiti, detti a *logica cablata*, che evolvevano in tanti modi diversi quante erano le istruzioni che essa era in grado di svolgere



# Logica Microprogrammata

- Nei microprocessori ad ogni istruzione corrisponde una sequenza di microistruzioni conservate in una memoria interna alla UC
  - La sequenza di microistruzioni ha il compito di generare le abilitazioni necessarie alla attuazione della istruzione
  - A tal fine un circuito interno alla UC provvede alla generazione di indirizzi per individuare una dopo l'altra le microistruzioni che un decodificatore trasforma in segnali di abilitazione
  - L'istruzione nel registro IR determina la posizione della prima microistruzione



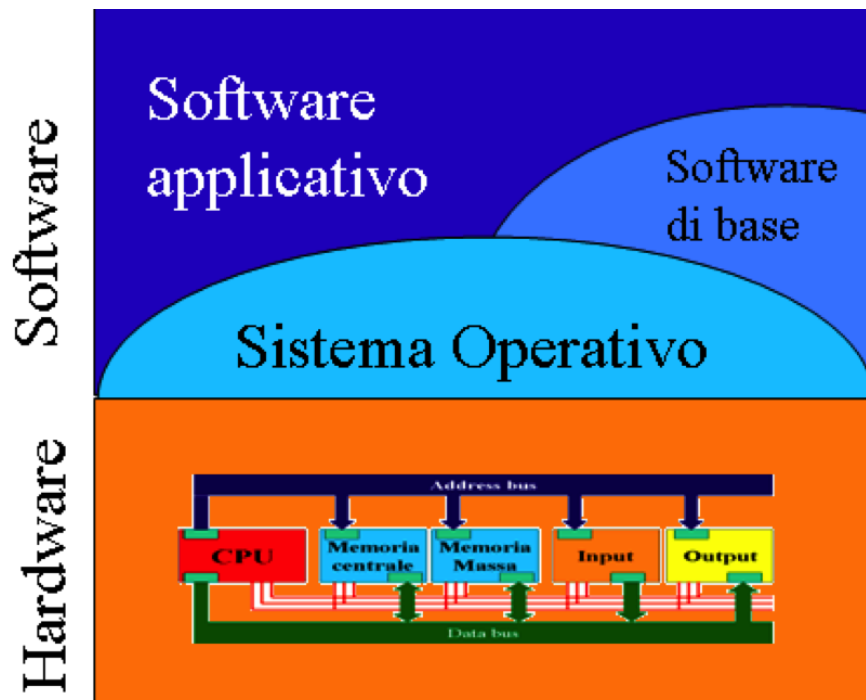
# Firmware e Software

- L'insieme dei micro programmi composti dalle microistruzioni memorizzate nella memoria interna alla UC prende il nome di *firmware*
- L'insieme di tutte le applicazioni del computer, quindi di tutti i programmi per computer, prende il nome di *software*.
  - In una accezione più ampia il termine software può essere inteso come tutto quanto può essere preteso dall'hardware: basta infatti inserire in memoria un programma diverso perché il sistema cambi le sue attività
  - Tra tutte le macchine automatiche il computer è un sistema polifunzionale in quanto può eseguire infinite funzioni sempre che venga progettato un programma per ogni applicazione



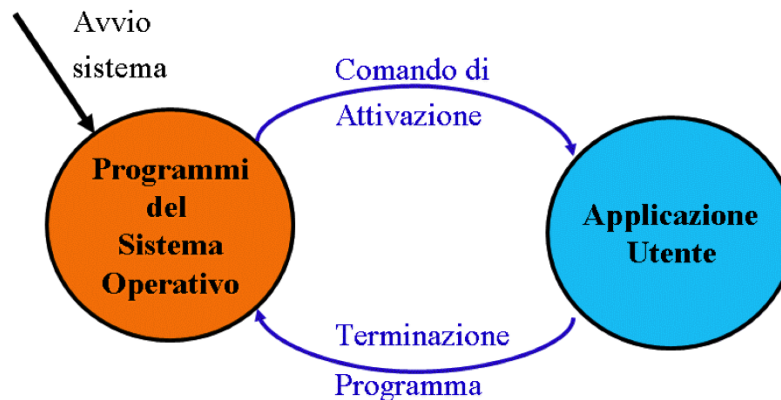
# Diversi tipi di Software

- Programmi che servono a tutti gli utenti del sistema
  - Classificati come *software di base*
    - sistemi operativi e traduttori dei linguaggi di programmazione
- Programmi che risolvono problemi specifici
  - *software applicativo*



# Sistema Operativo

- Il sistema operativo è un insieme di programmi che deve garantire la gestione delle risorse hardware in modo semplice ed efficiente a tutti gli utenti del sistema
  - persone oppure altre applicazioni
  - I primi calcolatori non avevano il sistema operativo
    - In essi il programmatore doveva prevedere tutto
    - Al termine dell'esecuzione del programma il programmatore o l'operatore del sistema doveva provvedere ad un nuovo caricamento in memoria ed ad una successiva attivazione
  - Con il sistema operativo il passaggio da una applicazione ad un'altra è svolto in automatico
    - La CPU si trova così ad eseguire i programmi del sistema operativo in alternanza con quelli applicativi



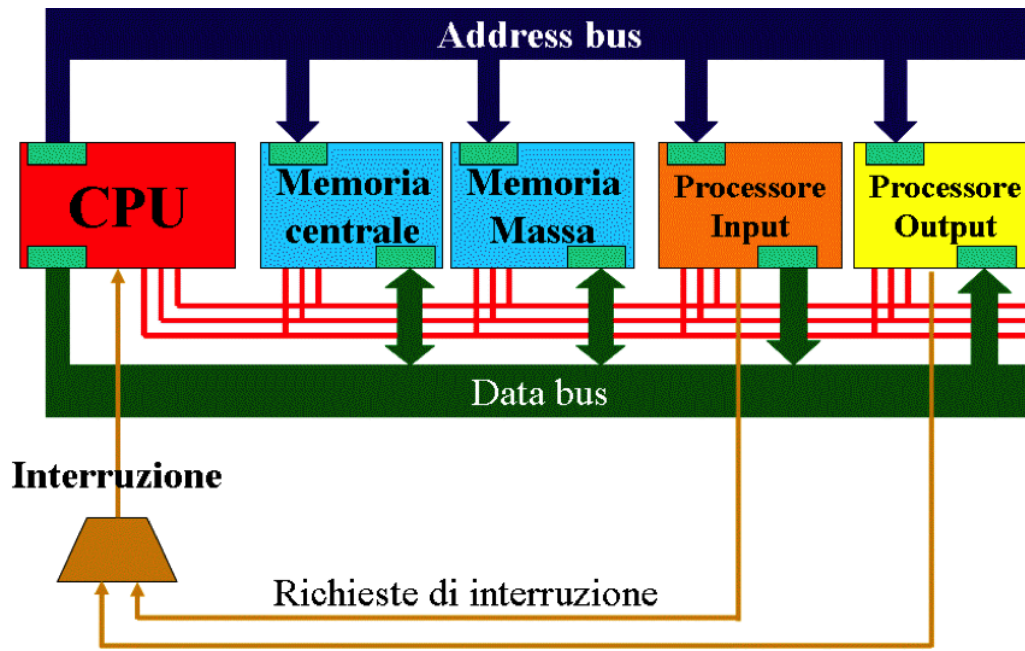
# Evoluzione del modello di Von Neumann: interruzioni

- Nel modello di Von Neumann, non era possibile sovrapporre i tempi delle operazioni di input con quelli dell'output
- *Introduzione di sistemi dedicati (canali)* il cui compito è scaricare la CPU della gestione di attività specifiche
  - I canali, con la loro autonomia possono lavorare anche contemporaneamente con la CPU
    - Es. canali di input ed output, processori dedicati alla grafica, alle operazioni sui numeri reali, alla acquisizione di segnali analogici
- Per rendere indipendenti i processori dedicati è stato introdotto nell'architettura hardware un segnale detto delle interruzioni mediante il quale una qualsiasi entità esterna alla CPU può richiederle attenzione

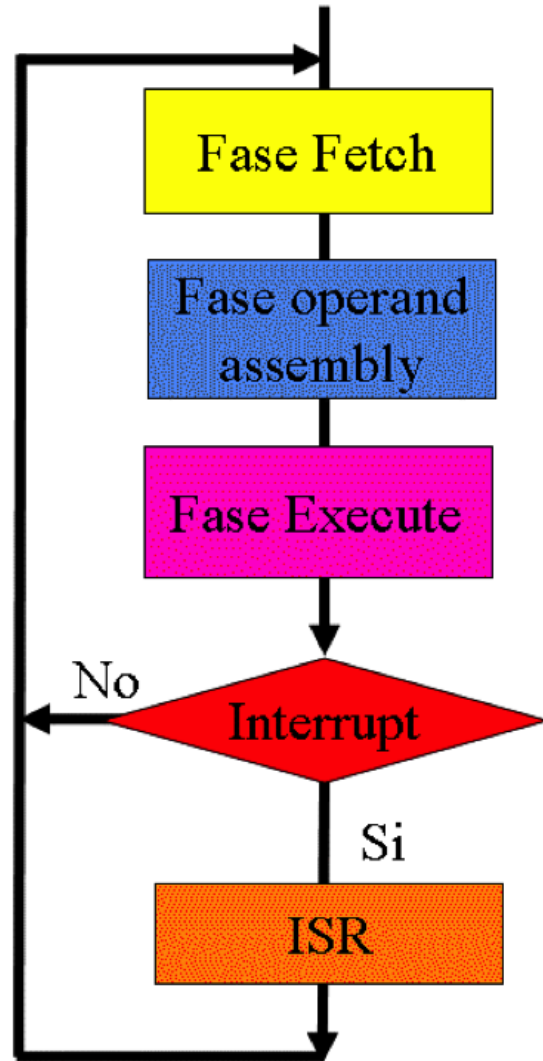


# Le interruzioni

- Con la presenza del *segnale di interruzione* la CPU può attivare un processore periferico e disinteressarsi delle sue attività
  - Quando un processore dedicato termina il suo compito, avanza una richiesta di interruzione al processore centrale e aspetta che gli venga rivolta attenzione
  - Mentre i processori periferici lavorano, la CPU può lavorare anch'essa a meno che non sia indispensabile quanto richiesto allo specifico processore



# ISR



- Per consentire alla UC di accorgersi del verificarsi di una interruzione il registro di condizione CC è stato dotato di un bit che diventa uguale ad uno quando arriva una interruzione
- La UC controlla il bit al termine delle esecuzione di ogni istruzione
  - Se è uguale zero procede normalmente con il prelievo dell'istruzione successiva
  - In caso contrario comincia l'esecuzione di un programma del sistema operativo, detto ISR (interrupt service routine) che ha come compito primario di capire la causa della interruzione, ossia quale dispositivo ha avanzato la richiesta
  - Nel caso si accorga della presenza di più richieste stabilisce quale servire per prima secondo criteri di importanza o priorità di intervento

# Evoluzione del modello di Von Neumann: le Cache

- Per ridurre i tempi di trasferimento dalla memoria centrale ai registri interni della CPU, viene replicata una porzione di memoria e posta tra memoria e CPU stessa. Tale memoria, molto veloce, viene chiamata *cache* e fa da buffer per il prelievo di informazioni dalla memoria centrale
  - Con operazioni particolari, istruzioni e dati vengono trasferiti dalla memoria centrale nella cache secondo la capacità di quest'ultima
    - La UC procede nelle tre fasi del suo ciclo al prelievo di istruzioni e operandi dalla cache
    - Quando la UC si accorge che il prelievo non può avvenire scatta un nuovo travaso dalla memoria centrale.
  - Se la cache è interna alla CPU viene detta di primo livello (L1)
  - Le cache di secondo livello (L2) sono invece esterne e solitamente un pò più lente di quelle di primo livello ma sempre più veloci della memoria centrale
    - La cache L2 risulta 4 o 5 volte più lenta della cache L1 mentre la RAM lo è addirittura 20 o 30 volte. I due livelli possono coesistere

# Gerarchie di Memorie

- Consente di offrire ai programmi l'illusione di avere una memoria grande e veloce
  - Nella gerarchia i livelli più prossimi alla CPU sono anche quelli più veloci, ma sono anche quelli con dimensioni più piccole visto il loro elevato costo
  - I livelli più lontani sono quelli che mostrano una capacità massima ed anche tempi di accesso maggiori

