

Appunti di Informatica

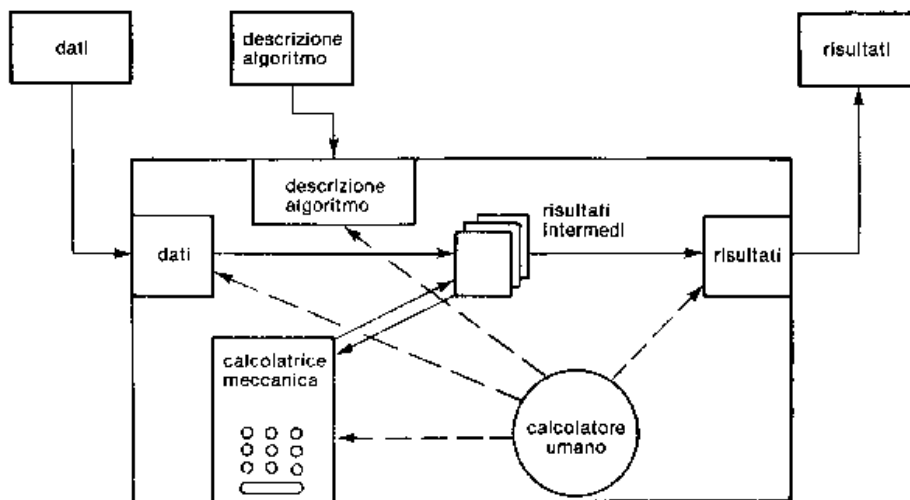
Struttura generale di un calcolatore

<i>Generalità sul processo di elaborazione automatica</i>	1
<i>Introduzione ed estrazione dei dati</i>	6
Generalità sull'introduzione e l'estrazione dei dati.....	6
<i>La scheda perforata</i>	7
<i>Il nastro perforato</i>	11
<i>Lettura di tipo magnetico</i>	13
<i>Lettura di tipo ottico</i>	15
<i>Codici a barre</i>	17
<i>Apparecchiature terminali: tastiere e unità video</i>	19
<i>Stampanti</i>	21
Stampanti seriali ad impatto.....	22
Stampanti parallele ad impatto	22
Stampanti non ad impatto	23
<i>Registratori grafici o "plotter"</i>	24
MEMORIZZAZIONE DEI DATI	25
Generalità	25
<i>Memorie a nuclei</i>	27
Scrittura di un nucleo magnetico	30
Lettura di un nucleo magnetico	31
<i>Memorie a semiconduttori</i>	34
<i>Memorie ausiliarie o di massa</i>	35
Nastri magnetici.....	35
Dischi magnetici.....	36
<i>Supporti magnetici di basso costo</i>	38
<i>Altri tipi di memorie</i>	39
FUNZIONI DI CALCOLO E DI GOVERNO.....	40
<i>Le istruzioni</i>	40
<i>L'unità di governo</i>	43
Funzionamento dell'unità di governo.....	44
<i>L'unità aritmetica e logica</i>	47
<i>Cenni sulle architetture dei sistemi di elaborazione</i>	48

Generalità sul processo di elaborazione automatica

Prima ancora di descrivere come avviene l'**elaborazione automatica dei dati** in un *calcolatore elettronico*, analizziamo come si svolge una elaborazione non automatica, meglio detta **manuale**, degli stessi dati. Consideriamo perciò il comportamento di una persona che deve risolvere un certo *problema*, applicando un *algoritmo* (il cui significato può anche esserle ignoto) ad un *insieme di dati*.

Supponiamo che questa persona abbia a disposizione dei fogli di carta, una matita e una calcolatrice meccanica tradizionale, e supponiamo che riceva dall'esterno un foglio sul quale sono riportati i *dati iniziali* ed un foglio sul quale invece è riportata la descrizione dell'*algoritmo* da applicare (sarebbe l'elenco delle *istruzioni* che definiscono la successione delle operazioni da effettuare sui *dati*).



Schema del processo di elaborazione manuale

Analizziamo il *flusso delle informazioni* durante l'esecuzione dell'algoritmo: esse, una volta introdotte nel sistema, passano per la calcolatrice meccanica, vengono elaborate, vengono trascritte una volta effettuata l'elaborazione e vengono infine portate all'esterno del sistema. Il tutto sotto il controllo del *calcolatore umano*, il quale supervisiona ogni singola fase del processo.

In questa visione, si individua subito ciò che viene chiamato **sistema di elaborazione**: si tratta del sistema costituito dal calcolatore umano, dalla calcolatrice meccanica e dai fogli di carta sui quali vengono trascritti (a mano nel nostro caso) i dati iniziali, l'algoritmo e i risultati finali.

Descriviamo adesso velocemente le principali funzioni elementari che si evidenziano dallo schema citato:

- a) **funzione di introduzione**: vengono introdotti nel sistema di elaborazione i *dati iniziali* del problema (forniti all'esterno), sui quali il calcolatore dovrà effettuare le operazioni previste dall'algoritmo (la cui descrizione è anch'essa fornita dall'esterno). Queste introduzioni vengono effettuate, nel nostro esempio, trascrivendo i dati e la descrizione dell'algoritmo dai fogli

provenienti dall'esterno sui fogli che fanno invece parte del sistema, oppure, più semplicemente, spostando i fogli dall'esterno all'interno del sistema;

- b) **funzione di estrazione:** questa funzione consiste nella trascrizione dei *risultati* dai fogli che fanno parte del sistema su un foglio che viene consegnato all'esterno (sarebbe la parte finale del processo);
- c) **funzione di memorizzazione:** questa funzione viene svolta dai fogli di carta sui quali il calcolatore umano riporta i dati manualmente. Questi fogli contengono sia i dati iniziali sia la descrizione dell'algoritmo sia gli eventuali valori intermedi sia i risultati finali;
- d) **funzione di calcolo:** l'algoritmo introdotto può richiedere l'esecuzione di varie *operazioni* sui dati, delle quali alcune possono essere di tipo *aritmetico* e altre di tipo *logico*. Le prime vengono eseguite utilizzando la calcolatrice meccanica e riportando i risultati sui fogli di carta che svolgono la funzione di memoria. Le operazioni di tipo logico vengono invece eseguite dal calcolatore umano, ma si può facilmente supporre l'esistenza di un ulteriore dispositivo autonomo in grado di svolgere analoghe funzioni;
- e) **funzioni di governo:** queste funzioni vengono svolte solo dal calcolatore umano; esso legge ed interpreta in modo corretto le istruzioni contenute nell'algoritmo, determina (in base ai risultati dei confronti) la sequenza secondo la quale tali istruzioni devono essere eseguite, governa e controlla il funzionamento della calcolatrice meccanica, gestisce i fogli di carta nei quali sono memorizzati i dati intermedi, provvede all'estrazione dei risultati ottenuti.

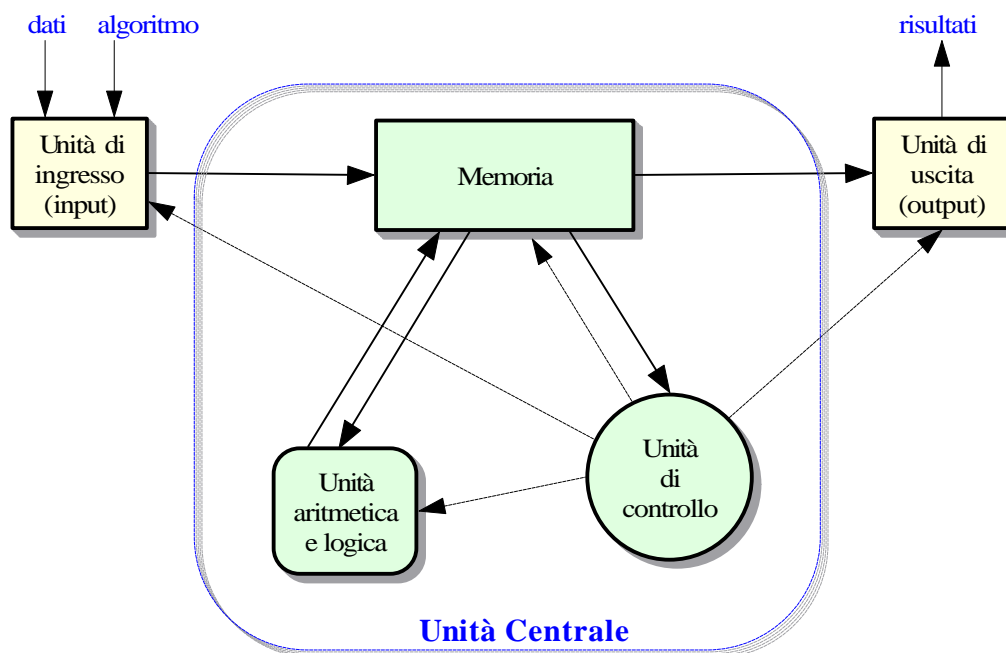
La prima cosa che risulta evidente da questa descrizione è che tutte le operazioni svolte dal calcolatore umano, incluse quelle di governo e controllo, possono essere facilmente automatizzate, ossia è facile pensare alla loro esecuzione da parte di un sistema completamente automatico, ad esempio un **calcolatore automatico**. Vediamo allora le principali differenze tra il sistema appena descritto, che potremmo definire "semi-automatico" per la presenza della calcolatrice meccanica, e un sistema completamente "automatico":

- nel sistema automatico, i fogli di carta utilizzati per registrare le varie informazioni vengono sostituiti da una **memoria**, ossia da un organo in

grado di ricevere, conservare e restituire in qualsiasi momento dei dati opportunamente codificati;

- la calcolatrice meccanica viene sostituita da un organo chiamato **unità aritmetica e logica**, capace di svolgere la funzione di calcolo sia aritmetico sia logico;
- il calcolatore umano viene infine rimpiazzato da un organo, detto **unità di governo**, che interpreta le istruzioni contenute nell' algoritmo da eseguire e controlla il funzionamento dell'intero sistema;
- le funzioni di introduzione e di estrazione dei dati vengono svolte da due differenti unità, dette rispettivamente **unità di introduzione** (o di **input**) e **unità di estrazione** (o di **output**).

L'insieme costituito dalla memoria, dall'unità aritmetica e logica e dall'unità di governo viene chiamato **CPU**, abbreviazione di *Central Processing Unit*.



Schema del processo di elaborazione automatica

Naturalmente, come nel caso del sistema "semi-automatico", l'unità di controllo svolge funzioni appunto di governo e controllo nei confronti dell'unità di ingresso, dell'unità aritmetica, della memoria e dell'unità di uscita. I dati e la descrizione dell'algoritmo vengono introdotti nell'unità centrale, attraverso l'unità di input (la

tastiera o un terminale), e vengono registrati direttamente nella memoria del calcolatore; i dati vengono poi prelevati dalla memoria, sempre sotto il controllo dell'unità di controllo, e inviati all'unità aritmetica e logica; questa elabora i dati tramite l'algoritmo e li restituisce alla memoria, che li registra nuovamente; infine, i dati finali, ossia i risultati ottenuti tramite l'unità aritmetica e logica, vengono passati dalla memoria all'unità di output (il video o la stampante) che provvede alla loro estrazione dal sistema.

Nella maggior parte dei casi, un calcolatore è dotato di due **livelli di memoria**: la **memoria principale** è in grado di immagazzinare una quantità limitata di dati e di renderli accessibili in tempi molto brevi, mentre le cosiddette **memorie ausiliarie** (o *memorie di massa*), sono in grado di immagazzinare grandi quantità di dati, il cui accesso richiede però un tempo relativamente più lungo.

Anche le unità di input o di output possono essere molteplici: l'insieme di tali unità e delle memorie ausiliarie prende il nome di **periferia**, in contrapposizione evidentemente all'unità centrale.

L'**architettura** di un calcolatore elettronico, ossia la scelta e la disposizione delle unità periferiche, viene realizzata in modi diversi, anche in relazione alla fascia di costi e di prestazioni verso le quali i costruttori intendono orientarsi. Dal punto di vista dei costi e delle prestazioni, i calcolatori possono essere infatti classificati nelle seguenti categorie fondamentali:

- **personal computer** (o *microcalcolatori*), che costano poco (tra 1 milione e qualche decina) e forniscono comunque ottime prestazioni per uso di tipo "hobbistico";
- **minicalcolatori** (o *minicomputer*), più costosi (si arriva a qualche centinaio di milioni) e più sofisticati, utilizzati soprattutto per il controllo dei processi aziendali;
- infine i **medi e grandi calcolatori**, costosissimi ma con prestazioni quasi inimmaginabili.

E' possibile inoltre collegare calcolatori di varie dimensioni mediante *linee di comunicazione* particolari in modo da formare le cosiddette **reti di calcolatori**.

Introduzione ed estrazione dei dati

Generalità sull'introduzione e l'estrazione dei dati

I dati da elaborare possono essere forniti ad un calcolatore mediante l'uso di supporti e di apparecchiature molto diversificati. A questo proposito, la prima cosa da sottolineare è che la correttezza di qualsiasi elaborazione svolta da un calcolatore sia condizionata alla correttezza dei dati che vengono introdotti: è perciò necessario sia che i dati posti in input siano esatti sia che l'**operazione di input** venga eseguita senza errori. E' infatti per evitare inconvenienti che le unità di input sono sempre dotate di dispositivi che controllano l'esattezza dell'operazione di lettura e permettono inoltre all'utente stesso di verificare che l'introduzione dei dati sia stata corretta.

Prima di cominciare la rassegna di un certo numero di unità input, è bene osservare che i dati possono essere forniti ad un calcolatore in diversi modi:

- i dati possono trovarsi su dei supporti che il calcolatore deve trattare in modo più o meno automatico e, naturalmente, sono registrati su quel supporto in modo tale che il calcolatore possa riconoscerli;
- i dati possono essere digitati sulle tastiere di apparecchiature collegate con il calcolatore oppure sulle tastiere di apparecchiature remote, per poi essere inviati al calcolatore mediante linee di comunicazione;
- i dati possono anche provenire da strumenti che misurano delle grandezze fisiche e che trasmettono poi le misure effettuate tramite segnali digitali riconoscibili dal calcolatore;
- i dati possono infine provenire da altri calcolatori collegati mediante linee di comunicazione.

Allo stato attuale, non sono ancora entrate nell'uso corrente apparecchiature in grado di riconoscere messaggi vocali di una certa complessità, come anche non si è ancora riusciti a costruire apparecchi che forniscano una lettura assolutamente affidabile di documenti scritti manualmente.

Per quanto riguarda invece l'**estrazione di dati**, si distinguono anche qui vari casi:

- i dati in uscita dal calcolatore vengono stampati su fogli di carta o visualizzati su uno schermo video in forma leggibile o in forma grafica su di un tracciatore di disegni (detto plotter) o ancora trasformati in segnali vocali;
- i dati in uscita possono essere registrati, con opportuni metodi di codifica, su supporti in grado di essere trattati automaticamente e perciò utilizzati nuovamente per eventuali altre elaborazioni;
- i dati in uscita possono essere trasmessi ad altri calcolatori o ad apparecchiature terminali mediante linee di comunicazione;
- i dati in uscita possono essere infine inviati a degli strumenti che permettono di controllare impianti industriali o apparecchi scientifici.

La scheda perforata

Questo è stato per lungo tempo il principale supporto per l'introduzione di dati in un calcolatore. Si tratta di un cartoncino rettangolare avente dimensioni standard (lunghezza mm 187.3 e larghezza mm 82.5) con particolari caratteristiche fisiche (resistenza alle sollecitazioni meccaniche, isolamento elettrico, ...). Spesso, uno degli angoli della scheda è smussato al fine di poter controllare visivamente l'esatta posizione di tutte le schede in un pacco.

La **scheda** viene suddivisa in un numero fisso di **colonne** verticali (80) e di **righe** orizzontali (12). Le colonne sono numerate in ordine crescente da 1 ad 80 da sinistra a destra ; delle righe, invece, sono numerate solo le 10 inferiori (che vanno perciò dalla riga 0 alla riga 9 partendo dall'alto); le altre due righe, chiamate solo convenzionalmente 11 e 12, si trovano in uno spazio libero sopra la riga 0 in posizione simmetrica delle righe 1 e 2 rispetto alla stessa riga 0. Le righe da 0 a 9 vengono dette **normali** mentre le righe 11 e 12 vengono dette **sopraelevate**.

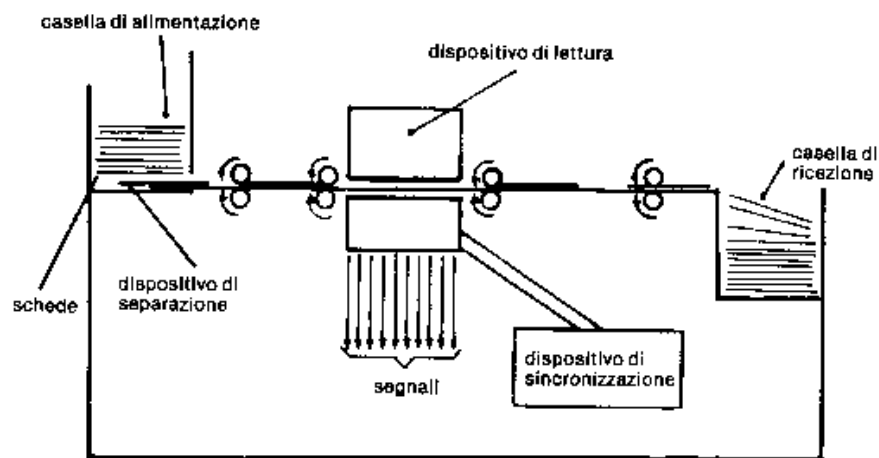
In ciascuna colonna della scheda può essere registrato un carattere mediante la perforazione di 1, 2, 3 o 4 fori rettangolari in posizione (cioè in righe) opportune. Le tabelle che forniscono la corrispondenza tra i fori rettangolari (o perforazioni) su una scheda e i simboli dell'alfabeto costituiscono il cosiddetto **codice di perforazione**.

La registrazione di un numero naturale compreso tra 0 e 9 può essere effettuata mediante la semplice perforazione della riga che porta il nome della cifra da

rappresentare: se si vuole, ad esempio, indicare la cifra 5, basta perforare, una sola volta, la riga numero 5.

La registrazione di una lettera dell'alfabeto o di altri simboli particolari (segni di interpunzione, operatori matematici, ...) viene effettuata mediante 2 o più perforazioni praticate nella stessa colonna.

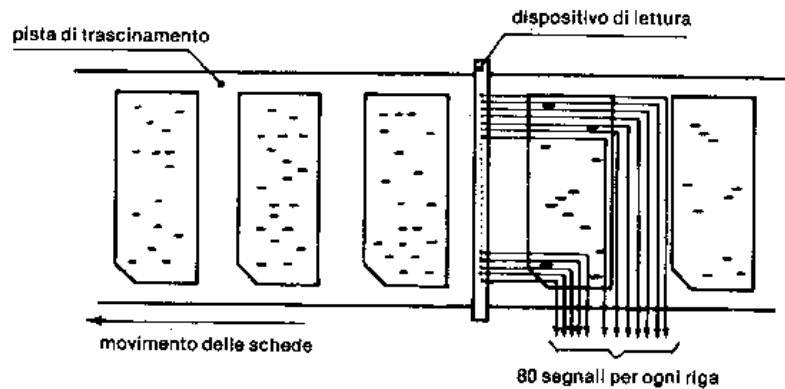
Le schede vengono lette in modo automatico dai cosiddetti **lettori di schede**: ce ne sono di diversi tipi a seconda della velocità di funzionamento e del modo con cui effettuano la lettura. Le velocità di funzionamento vanno da circa 100 schede al minuto per i lettori più lenti a più di 3000 schede al minuto per i più veloci.



Schema di un lettore di schede

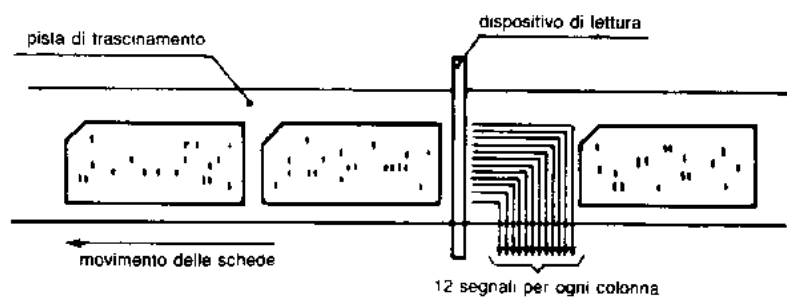
I sistemi utilizzati per l'avanzamento delle schede nel lettore sono di due tipi, il che significa che si hanno fondamentalmente due tipi di lettori:

- nei lettori con **funzionamento parallelo**, la scheda viene trascinata in direzione perpendicolare al suo lato più lungo: perciò, il dispositivo legge una riga alla volta, trasmettendo quindi contemporaneamente 80 segnali (cioè il numero delle colonne) in corrispondenza di ogni riga; questo significa che i caratteri perforati su di una scheda vengono riconosciuti per intero solo dopo la lettura dell'ultima riga;



Letture di una scheda in modo parallelo

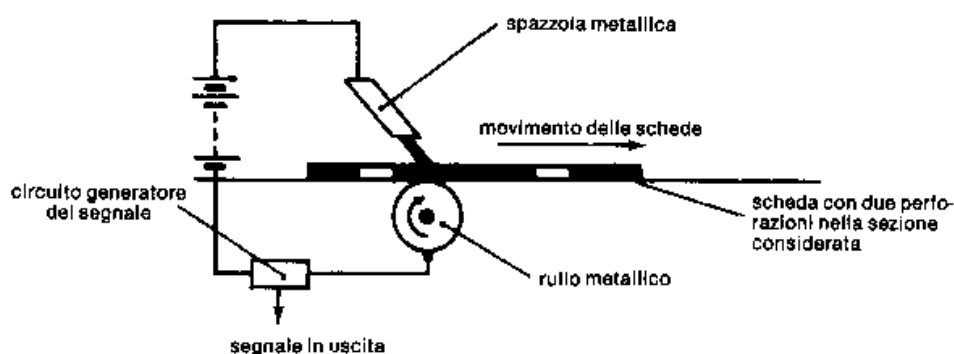
- nei lettori con **funzionamento seriale**, la scheda viene invece trascinata all'interno del lettore in direzione perpendicolare al suo lato più corto. Allora, il dispositivo di lettura legge la scheda una colonna alla volta, trasmettendo quindi 12 segnali (cioè il numero delle righe) in corrispondenza ad ogni colonna della scheda. Questo sistema di lettura, al contrario del precedente, permette di riconoscere i caratteri perforati su una scheda 1 alla volta.



Letture di una scheda in modo seriale

Il funzionamento di tipo seriale è solitamente utilizzato per apparecchiature di basso costo rispetto a quelle funzionanti in parallelo; anche per questo sono più lente.

La lettura vera e propria delle schede può essere realizzata mediante dispositivi di tipi *elettromeccanico* o di tipo *fotoelettrico*. Un **dispositivo elettromeccanico** è costituito da un certo numero di spazzolini metallici che sono in contatto con un rullo metallico; questo ruota attorno al proprio asse ed è sottoposto ad una certa tensione.



Letture di tipo elettromeccanico

La scheda passa allora tra il rullo e gli spazzolini: in corrispondenza di ogni perforazione si chiude il circuito elettrico costituito dallo spazzolino interessato e dal rullo, il che genera un impulso elettrico che il calcolatore va ad interpretare. Questo tipo di lettura presenta dei grossi limiti di velocità: difatti, oltre certe velocità, trattandosi appunto di un dispositivo meccanico, bisogna fare i conti con problemi di inerzia meccanica degli spazzolini che non rendono affidabile la lettura. Ed è questo il motivo per cui, per i sistemi di lettura ad alte velocità, si preferisce utilizzare i **lettori di tipo fotoelettrico**: in questo caso, la scheda passa attraverso una sorgente luminosa e alcune cellule fotosensibili; in presenza di una perforazione, il raggio emesso dalla sorgente colpisce la corrispondente cellula fotoelettrica che genera l'impulso elettrico che il calcolatore deve interpretare.

Anche per quanto riguarda le apparecchiature per la **perforazione delle schede**, cioè per la registrazione dei dati, si distinguono due tipi fondamentali:

- ci sono i cosiddetti **perforatori di schede**, che realizzano le perforazioni in modo automatico: questi sistemi servono generalmente per effettuare l'estrazione dei dati dal calcolatore registrandoli su schede che possono essere nuovamente trattate in modo automatico;
- ci sono poi le apparecchiature per la perforazione manuale delle schede e che vengono dette **perforatrici**: esse comprendono una tastiera, una casella di alimentazione, un sistema di trascinarsi, un dispositivo di perforazione ed una casella di ricezione. Premendo il tasto corrispondente al carattere desiderato, vengono effettuate le dovute perforazioni su una colonna della scheda.

L'esatta perforazione delle schede può essere controllata mediante una apparecchiatura detta **verificatrice**: essa è molto simile ad una perforatrice, con l'unica differenza che al posto dei punzoni ci sono degli spazzolini che verificano che le perforazioni siano state corrette, ossia che siano state fatte nelle giuste posizioni.

Per quanto riguarda il trattamento automatico delle schede, sono stati elaborati nuovi sistemi detti **sistemi a schede perforate**; un *sistema a schede perforate* comprende, in genere, i seguenti dispositivi:

- una **selezionatrice di schede** che ordina le schede in base ad un codice perforato su una o più colonne;
- una **inseritrice comparatrice**, che mette insieme 2 o più gruppi di schede sempre in base ad un criterio indicato sulle schede stesse;
- una **riproduttrice**, che perfora sulle schede dei dati fissi precedentemente impostati e, allo stesso tempo, controlla la correttezza delle perforazioni;
- una **tabulatrice**, che costituisce la parte fondamentale di un sistema a schede perforate e svolge molteplici funzioni: legge i dati perforati sulle schede, li elabora in base a dei codici prestabiliti ed infine stampa su prospetti (o tabulati) i dati letti e gli eventuali risultati calcolati;
- infine, una **calcolatrice** che esegue i calcoli per conto della tabulatrice.

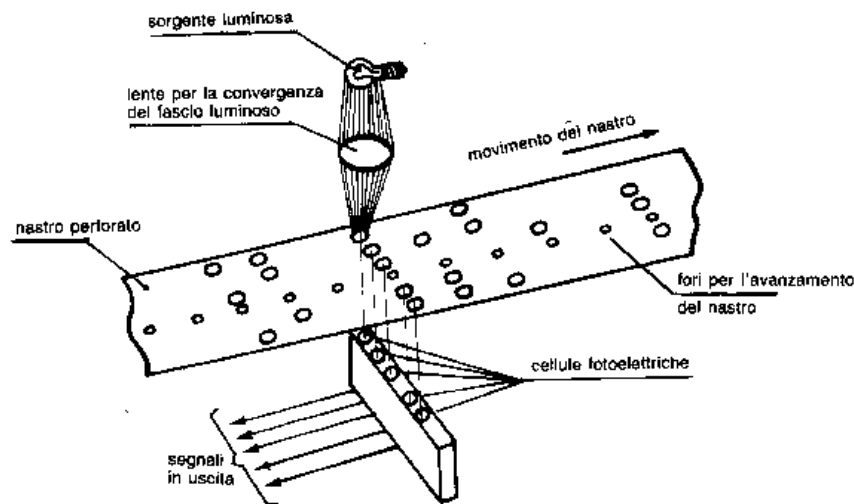
Il nastro perforato

Un altro supporto spesso usato per l'introduzione dei dati nei calcolatori elettronici è stato il **nastro perforato**. Si tratta di una striscia continua (detta *banda*) di carta opaca, di lunghezza variabile, normalmente avvolta su bobine, sulla quale i dati vengono registrati mediante perforazioni disposte su righe perpendicolari all'asse. Il nastro può essere considerato come suddiviso in tante righe, ciascuna delle quali rappresenta un carattere in base alle posizioni e al numero delle perforazioni presenti sulla riga considerata.

I vari tipi di nastro perforato possono distinguersi sia per la forma dei fori (circolari o rettangolari) sia per il numero massimo delle perforazioni presenti su

una stessa perpendicolare all'asse del nastro, sia per il codice di perforazione usato.

Una apparecchiatura per la lettura del nastro perforato può essere così schematizzata:



Letture di tipo fotoelettrico

C'è per prima cosa una *bobina di alimentazione* attorno alla quale è avvolto il nastro da leggere; su comando del *dispositivo di avanzamento*, la bobina comincia a ruotare svolgendo il nastro, il quale si avvia verso la *stazione di lettura*; passato sotto tale stazione ed effettuata la lettura, il nastro viene infine riavvolto su un'altra bobina, detta *bobina di ricezione*, che, naturalmente, all'inizio era vuota.

La lettura del nastro perforato può essere effettuata, come per le schede perforate, da dispositivi *elettromagnetici* o *fotoelettrici*. Anche qui, quelli del primo tipo sono poco utilizzati, sia per il fatto che usurano facilmente il nastro sia per i già citati motivi di lentezza. Più affidabili ed efficienti sono i **dispositivi fotoelettrici**, che consentono la lettura di 1000 caratteri al secondo .

Le apparecchiature che servono per la perforazione automatica dei dati in uscita da un calcolatore sul nastro perforato hanno uno schema perfettamente analogo a quello dei lettori: è chiaro che il dispositivo di lettura è sostituito da uno di perforazione che, in corrispondenza dei bit 1 (che segnalano la presenza di un carattere), comanda elettromagneticamente un punzone che effettua la perforazione voluta.

Da notare che sia la lettura sia la perforazione di un nastro avvengono in modo **seriale** (ossia in direzione perpendicolare al lato più corto del nastro), dato che il

nastro si svolge con continuità dalla bobina di alimentazione per riavvolgersi poi su quella di ricezione.

Il numero massimo delle perforazioni che possono essere effettuate su ogni riga (ossia il numero di tracce del nastro) determina il numero dei caratteri che si possono rappresentare: perciò sui nastri a 5 tracce si possono rappresentare $2^5=32$ caratteri diversi, su quelli a 6 tracce $2^6=64$ caratteri, su quelli a 7 tracce $2^7=128$ e così via. Tuttavia, nei nastri a 8 tracce, dove teoricamente si potrebbero rappresentare $2^8=256$ caratteri, si preferisce riservare una traccia a scopo di controllo, per cui i caratteri effettivamente rappresentabili sono solo 128, come i nastri a 7 tracce. Questa traccia utilizzata per controllo si chiama **traccia di parità** (oppure **di disparità**): il suo scopo è quello di rendere sempre pari (o sempre dispari) il numero delle perforazioni relative ad ogni carattere. Questo permette una facile rilevazione di eventuali errori di perforazione o di lettura da parte degli appositi dispositivi.

Lettura di tipo magnetico

L'impiego delle schede perforate come supporto sul quale registrare le informazioni da inviare ad un calcolatore richiede l'esecuzione di una particolare operazione, ossia la **conversione dei supporti**: infatti, i dati che devono essere inviati al calcolatore si trovano originariamente su determinati supporti (ad esempio fogli di carta con tracciati particolari) e devono perciò essere perforati manualmente, cioè *convertiti*, sulla scheda in modo da poter essere trattati automaticamente. Questa operazione di conversione dei supporti è piuttosto complessa, in quanto, oltre ad essere particolarmente lunga, deve essere effettuata senza errori se si vuole che il calcolatore compia correttamente le proprie funzioni. Per poter ovviare al problema, ossia per evitare eventuali errori di perforazione, ci si è sforzati di costruire delle apparecchiature che fossero in grado di trattare i supporti originali e di leggere ed interpretare i dati espressi su di essi mediante gli stessi simboli grafici utilizzati dall'uomo. In altre parole, è stata eliminata la fase della *perforazione manuale* e si è invece provveduto ad un *trattamento automatico* anche dei supporti originari dai quali vanno estratti i dati. Il trattamento di tali supporti originari, ossia la loro lettura, si fonda su varie possibilità: innanzitutto la possibilità di "esplorare" la superficie di tali supporti al fine di rilevare su di essa la presenza di segni o caratteri convenzionali il più possibile simili ai segni grafici

tradizionali; una volta rilevata tale presenza, la possibilità di tradurla in impulsi elettrici; questi impulsi elettrici, a loro volta, non sono altro che forme d'onda differenziate che il calcolatore può facilmente riconoscere come rappresentative di un particolare carattere.

I principali sistemi utilizzati per la compilazione e il successivo riconoscimento dei dati sui supporti originali sono i seguenti:

- il **sistema di lettura magnetica**, detto anche **MICR** (*Magnetic Ink Character Recognition*);
- il **sistema di lettura ottica**, detto anche **OCR** (*Optical Character Recognition*);
- il **sistema di lettura dei codici a barre**.

Nel *sistema di lettura ottica*, i caratteri vengono scritti sui supporti con forme ed inchiostri tradizionali (privi cioè di particolari proprietà); tuttavia il problema del loro riconoscimento, come vedremo, è piuttosto complesso. Nel *sistema di lettura magnetica*, invece, i caratteri vengono scritti sul supporto originale con un particolare inchiostro che contiene del materiale ferromagnetico. Difatti, questo sistema si basa sul seguente principio: i caratteri, essendo scritti con questo particolare materiale, nel momento in cui entrano in contatto con una **testina di lettura**, generano un particolare campo magnetico, che cambia da carattere a carattere; il sistema di lettura è in grado di rilevare i segnali elettrici prodotti dalla variazione nel tempo di tale campo magnetico. In ogni caso, per essere precisi, diciamo che il campo magnetico nasce prima che i caratteri passino sotto la testina di lettura: infatti, prima che ciò accada, i supporti originali vengono fatti passare sotto una **testina di magnetizzazione** (solitamente un magnete permanente): l'ossido di ferro contenuto nell'inchiostro si magnetizza e genera un campo magnetico strettamente legato alla forma del carattere scritto; a questo punto, come detto prima, il dispositivo di lettura è facilmente in grado di riconoscere il campo magnetico e di trasmettere al calcolatore il carattere corrispondente.

Il principale vantaggio di questo sistema proviene dal fatto che, se i caratteri scritti con l'inchiostro magnetico si trovassero accidentalmente ricoperti (totalmente o parzialmente) da macchie o segni estranei, la lettura automatica non ne verrebbe compromessa in quanto solo l'inchiostro speciale è in grado di

sviluppare il campo magnetico e quindi la lettura resta limitata al carattere che ha generato tale campo.

Dal punto di vista della **logica di riconoscimento**, i caratteri possono essere classificati in due fondamentali categorie: ci sono i **caratteri formali**, per i quali il riconoscimento è reso possibile dalla loro forma grafica, e i **caratteri codificati**, che sono invece identificabili solo dopo un esame delle parti che li compongono.

Esistono diversi tipi di caratteri utilizzati per la lettura magnetica: negli USA il tipo più diffuso è il cosiddetto **E13B**, che appartiene alla categoria dei caratteri formali; in Europa viene invece utilizzato un carattere magnetico di tipo codificato, detto **CMC7**: ogni carattere contiene all'interno della sua forma grafica un codice su cui è basato il riconoscimento da parte della macchina. Un carattere CMC7 è costituito da 7 barrette verticali, disposte in modo da formare 6 intervalli aventi 2 larghezze diverse: a seconda della loro posizione relativa, queste barrette determinano un codice incorporato direttamente nel carattere. Le 7 barrette sono inoltre di lunghezza tale da rappresentare il carattere con la sua forma usuale, anche per dare la possibilità di un facile riconoscimento visivo. Infine, il riconoscimento di questi caratteri avviene valutando le misure dei 6 intervalli e la loro combinazione.

I sistemi che funzionano con caratteri magnetici in generale sono impiegati generalmente nel campo bancario per il trattamento automatico degli assegni e delle cambiali.

Un normale **lettore di caratteri magnetici** si compone di un *dispositivo di alimentazione*, di una *pista di scorrimento per la lettura*, di un *dispositivo elettronico per il riconoscimento dei caratteri magnetici* e di due o più *caselle di ricezione*.

Lettura di tipo ottico

Il **sistema di lettura ottica** è un procedimento automatico per l'interpretazione di simboli grafici, stampati, aventi forma tradizionale. Le apparecchiature utilizzate per la lettura ottica possono trattare una notevole varietà di simboli grafici: alcuni lettori accettano caratteri di forma ben definita, altri invece accettano, anche mescolati, caratteri di forme diverse ed anche caratteri scritti a mano. Naturalmente, maggiore è la gamma dei caratteri che il calcolatore è in grado di riconoscere maggiore è il suo prezzo. L'obiettivo che si vuol raggiungere utilizzando la lettura ottica, così come con la lettura magnetica, è la possibilità di

utilizzare il documento originale direttamente come supporto per l'introduzione automatica dei dati nel calcolatore.

Le tecniche per effettuare la lettura ottica sono molteplici; la caratteristica comune a tutte è l'impiego di **dispositivi fotosensibili**, mediante i quali è possibile rilevare segnali ottici e trasformarli poi in impulsi elettrici.

Di seguito mostreremo le principali caratteristiche di soli quattro metodi per il rilevamento e il riconoscimento dei caratteri. Il primo di questi è il cosiddetto **metodo di rilevamento "a scansione"**: ogni carattere si trova in una ristretta zona rettangolare; si comincia a supporre che tale rettangolo sia ulteriormente suddiviso in una serie finita di quadrati più piccoli; tramite un sistema ottico, viene proiettato sulla faccia del documento un raggio luminoso di dimensioni finite; tale raggio colpisce ogni singolo rettangolo e scandisce tutti i quadratini minori di cui esso si compone; a questo punto, a seconda che il raggio luminoso colpisca una zona contenente più o meno inchiostro, viene diffusa più o meno luce; questa luce viene raccolta da un dispositivo (detto **fotomoltiplicatore**), che la converte, a seconda della sua intensità, nel corrispondente segnale elettrico.

Altro sistema di rilevamento è quello di tipo **analogico**: anche qui viene sfruttato il principio secondo cui, una volta che il raggio luminoso ha colpito una certa zona del documento, esso si riflette in modo diverso al variare della inchiostrazione. La differenza con il precedente sistema è che al carattere da riconoscere viene sovrapposta una rete rettangolare costituita da varie **cellule fotosensibili**: il carattere viene proiettato e messo a fuoco su tale rete in modo automatico e ogni fotocellula trasforma la quantità di luce che la colpisce in un segnale elettrico di ampiezza ad essa proporzionale. Si generano quindi una serie di impulsi elettrici i cui valori, opportunamente esaminati, permettono il riconoscimento del carattere in questione.

Un altro metodo importante è il **metodo di riconoscimento "a matrice bidimensionale"**: esso consiste essenzialmente nel confrontare una matrice bidimensionale di valori ricavati dal carattere in esame con dei *valori campione* già presenti nella memoria della macchina. La trasformazione del carattere in una matrice si ottiene nel modo seguente: preso il rettangolo che contiene il carattere, si considera il reticolo di piccoli elementi di cui a sua volta è costituito il rettangolo; in ognuno di tali elementi c'è una certa quantità di inchiostro: ad ogni elemento si associa un valore corrispondente alla quantità di inchiostro rilevata. Una volta ottenuta la matrice, la si confronta con le *matrici campione* che la macchina già possiede (mediante la semplice sovrapposizione); il carattere viene

infine individuato in base alla matrice campione che presenta la minima differenza con la matrice rilevata.

Infine, l'ultimo metodo di riconoscimento è quello detto **morfotopologico**: la matrice rappresentativa del carattere in esame viene analizzata allo scopo di individuare in essa la presenza di particolari elementi che descrivono la forma geometrica del carattere, a prescindere quindi dalle sue dimensioni e dallo stile con cui è scritto. La forma del carattere viene stabilita analizzando la curvatura dei tratti di inchiostro che lo compongono: per esempio, ci possono essere delle zone bianche interne o esterne al carattere, chiuse o aperte, di forma circolare, triangolare, rettangolare e così via. Ad esempio, il carattere 8 presenta due zone bianche chiuse circolari una sopra l'altra; oppure, il carattere 0 presenta una sola zona bianca, chiusa e circolare; ancora, il carattere 4 presenta una zona bianca, chiusa e di forma rettangolare. In questo modo, la matrice bidimensionale iniziale, rappresentativa del carattere, viene trasformata in un insieme più limitato di valori, che rappresentano le caratteristiche "geometriche" del carattere; tali valori vengono confrontati con i valori che rappresentano invece le caratteristiche della forma dei *caratteri campione*; alla fine, come nel caso precedente, la macchina opererà per il carattere campione i cui valori presentano la minima differenza con i valori del carattere rilevato. E' immediato dedurre che, con un metodo del genere, se la definizione dei caratteri campione è abbastanza generale, è possibile riconoscere caratteri appartenenti a stili molto diversi.

In ogni caso, tutti i sistemi di lettura ottica, ciascuno con i suoi pregi, presentano un grosso inconveniente comune: si tratta della inevitabile sensibilità del dispositivo di lettura ad eventuali *segni* presenti accidentalmente nella zona del carattere da analizzare, dovuti ad una stampa difettosa o alla manipolazione subita dai documenti stessi. Questi segni possono facilmente dar luogo ad una lettura difettosa e quindi a interpretazione errate dei caratteri. Da notare che invece, come si è sottolineato in precedenza, questo inconveniente non esiste per i sistemi di lettura magnetica, dove tali eventuali segni non compromettono la configurazione dei campi magnetici generati dai caratteri.

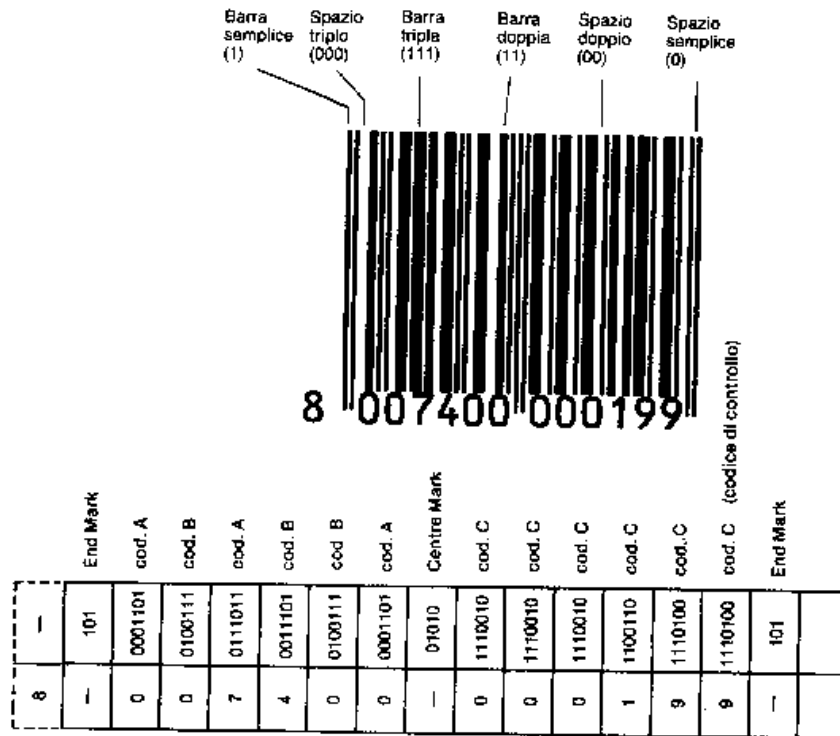
Codici a barre

Un metodo per l'input dei dati abbastanza diffuso soprattutto nelle attività commerciali è quello basato sulla marcatura delle confezioni dei prodotti mediante

codici a barre e sulla lettura di tali marcature mediante dispositivi più o meno automatici. Esistono vari tipi di codici a barre, ma tutti sono basati su successioni di barre scure separate da spazi vuoti chiari. Sia le barre sia gli spazi intermedi possono avere spessori diversi e questo allo scopo di rappresentare, mediante un codice binario, delle cifre decimali e in qualche caso anche lettere. Il codice più noto, che ci accingiamo a descrivere, è il **codice EAN** (*European Article Numbering*). Questo codice, derivato dall'americano **UPC** (*Universal Product Code*), nella sua forma denominata **EAN 13** permette di rappresentare 12 cifre decimali più una tredicesima espressa però in forma non codificata: queste 13 cifre permettono così di codificare la nazione di origine del prodotto, l'azienda produttrice e il prodotto stesso, oltre ad un codice di controllo. Il codice EAN 13 è *bidirezionale*, nel senso che le apparecchiature di lettura possono leggerlo sia da sinistra a destra sia da destra a sinistra.

Una marcatura in codice EAN 13 è dunque costituita da una sequenza di elementi base, chiari e scuri, tutti dello stesso spessore, che possono essere raggruppati fino ad un massimo di 4 per costituire barre e spazi più o meno larghi. Ogni elemento base viene interpretato come una cifra binaria: per cui 1 è l'elemento base scuro e 0 è quello chiaro; una intera marcatura è composta, da sinistra verso destra, da 95 elementi base che svolgono le seguenti funzioni:

- ci sono in primo luogo 3 elementi base fissi, precisamente 101, che costituiscono un primo *End Mark*;
- poi altri 42 elementi base che vengono interpretati come 6 cifre decimali (per cui ci sono 7 elementi base per ogni cifra);
- altri 5 elementi base fissi (che sono 01010) che costituiscono il *Centre Mark*;
- ancora 42 elementi base interpretati come altre 6 cifre decimali (sempre 7 elementi per ogni cifra). Da notare che l'ultima di queste 6 cifre decimali è una cifra di controllo che viene calcolata a partire dalle altre 5;
- infine, altri 3 elementi base fissi (101) che costituiscono il secondo *End Mark*.



Una marcatura in codice EAN 13

Gli elementi che costituiscono gli *End Mark* e il *Centre Mark* vengono stampati con una lunghezza maggiore degli altri; inoltre, tutte le cifre decimali vengono riprodotte anche in chiaro. Invece, la tredicesima cifra è stampata soltanto in chiaro sulla sinistra della marcatura. In definitiva, una marcatura EAN 13 contiene soltanto 12 cifre decimali codificate, mentre la tredicesima, non codificata, determina la codificazione delle altre dodici, che non è unica.

Apparecchiature terminali: tastiere e unità video

L'introduzione delle informazioni in un calcolatore viene realizzata in due fasi distinte: la prima consiste nella scrittura di queste informazioni sul supporto; la seconda nel trattamento automatico di tale supporto mediante delle apparecchiature che, una volta letti i dati, li inviano opportunamente codificati al calcolatore.

Esiste comunque la possibilità di fondere tali fasi, con conseguente velocizzazione del procedimento, in modo da introdurre dati nel calcolatore senza usare dei *supporti intermedi*. Le apparecchiature che permettono di fare questo si

chiamano **terminali**. Esse sono collegate direttamente, o anche tramite una *linea di trasmissione*, con il calcolatore e l'introduzione dei dati avviene mediante la loro battitura su apposite tastiere.

Il terminale può essere dotato di una apparecchiatura per la **stampa**, ma più spesso è collegato ad una **unità video**. Stampanti e video svolgono le funzioni di *unità di output*, ma visualizzano anche i dati introdotti tramite la tastiera.

Una unità video è in generale costituita da un *tubo a raggi catodici*, così come i normali televisori, anche se si trovano altri display realizzati con tecnologie più complesse.

Nelle unità video di *tipo televisivo*, la faccia interna dell'estremità del tubo catodico svolge le funzioni di *schermo* ed è ricoperta da uno strato di sostanze fotosensibili, dette *fosfori*, che vengono eccitate da *pennelli di elettroni*. L'emissione di luce da parte dei fosfori eccitati ha una durata infinitesima e quindi ogni punto dello schermo deve essere *rifrescato* 50 o 60 volte al secondo. Se tale frequenza dovesse risultare insufficiente, si avrebbe un tremolio dell'immagine chiamato tecnicamente *flickering*.

Nei **video monocromatici**, vengono impiegati dei fosfori che emettono luce bianca o verde o color ambra; nei **video a colori**, sono invece necessari fosfori che emettono luce blu, verde e rossa: una opportuna combinazione di questi tre colori base permette di ottenere qualsiasi gradazione della scala cromatica.

Se sullo schermo di una unità video possono essere stampati soltanto *caratteri alfanumerici*, allora il video si dice **alfanumerico**; invece, se è possibile visualizzare anche immagini, si parla di **video grafico**. Un video alfanumerico di carattere professionale permette di visualizzare 24 righe orizzontali di 80 caratteri ciascuna (cioè 80 colonne verticali), per un totale di 1920 caratteri. Ogni carattere può essere rappresentato in minuscolo e in maiuscolo ed è anche possibile ottenere effetti particolari, quali l'inversione (carattere scuro su fondo chiaro), l'intermittenza, la luminosità più intensa di una parola o di un carattere. I video grafici possono essere sia monocromatici sia a colori mentre i video alfanumerici sono in genere monocromatici. I video grafici, inoltre, permettono di controllare, mediante opportuni programmi, i singoli punti dello schermo, considerando cioè quest'ultimo come un reticolo composto da un numero rilevante di righe e di colonne. I punti del reticolo prendono il nome di **pixel** ed un video grafico di media capacità può essere visto come un reticolo di 480 righe e 640 colonne per un totale di 307200 pixel. Ovviamente, un video grafico può essere usato per visualizzare *testi alfanumerici*: anzi, le sue caratteristiche permettono una grande flessibilità di

rappresentazione e consentono di ottenere testi composti con caratteri di forma e dimensioni qualsiasi.

Le unità video, sia alfanumeriche sia grafiche, sono dotate di dispositivi che permettono all'utente di posizionarsi molto velocemente in qualsiasi punto dello schermo con il cosiddetto **cursore**: questo può essere spostato o mediante **tasti** o mediante una **tavoletta grafica** o mediante il **mouse** o infine mediante una **penna luminosa**. Ci sono anche schermi sensibili al contatto, i cosiddetti **touch-screen**.

I **terminali** sono talvolta fisicamente lontani dai calcolatori (sono detti perciò *terminali remoti*): in questo caso, il collegamento viene effettuato mediante cavo telefonico o mediante altre apparecchiature che effettuano la trasmissione dati. Invece, per *terminale di console* si intende un terminale collegato direttamente con il calcolatore usato.

Stampanti

Le **stampanti** sono unità di output che provvedono alla stampa dei risultati di determinate elaborazioni, mediante i simboli grafici abituali, su un supporto (generalmente continuo) di carta. Da un punto di vista prestazionale, le stampanti possono essere divise nelle seguenti categorie:

- **stampanti seriali**, che stampano solo un carattere per volta;
- **stampanti parallele**, che stampano contemporaneamente tutti i caratteri di una riga;
- **stampanti a pagina**, che stampano contemporaneamente una intera pagina.

Dal punto di vista, invece, della *tecnica di stampa*, possiamo così classificare le stampanti:

- **stampanti a impatto**, nelle quali si verifica un urto meccanico contro il nastro inchiostro e la carta;
- **stampanti non a impatto**, nelle quali l'impressione della carta avviene con tecniche fisiche o chimiche che non richiedono urti meccanici

Stampanti seriali ad impatto

In questo tipo di stampanti, il carattere può essere stampato o "per punti" o "ad immagine intera". Nel primo caso, un carattere viene formato da una serie di **aghi** che, comandati mediante elettromagneti, battono sul nastro inchiostro e quindi sulla carta, in modo da realizzare il carattere desiderato. Il carattere viene formato mediante un insieme di punti appartenenti ad una matrice di dimensioni variabili: per esempio 7 righe e 5 colonne per una matrice di 35 punti, oppure 9 righe e 7 colonne per una matrice a 63 punti. Il dispositivo di stampa contiene soltanto gli aghi necessari per stampare una colonna, per cui esso si muove lungo una riga di stampa a velocità costante. Alcuni modelli di **stampanti ad aghi** permettono sia di modificare il carattere di stampa (nella forma o nella grandezza) sia di aumentare la densità dei punti che compongono il carattere, il tutto per migliorare la qualità della stampa.

Nel caso, invece, di **stampa ad immagine intera**, i caratteri sono conati su appositi supporti, detti **testine**, che si posizionano lungo la riga in corrispondenza dei punti in cui devono essere stampati i caratteri. L'impressione avviene tramite l'impatto provocato da un elettromagnete. Le testine hanno solitamente la forma di sfere o cilindri o di margherite.

Le velocità ottenibili con stampanti ad impatto ad immagine intera non sono notevoli, dell'ordine di 30 caratteri al secondo. Migliori sono le prestazioni delle stampanti ad aghi, che arrivano fino a diverse centinaia di caratteri al secondo nei tipi migliori. Naturalmente, la velocità di stampa diminuisce quando è elevata la qualità di stampa che si è fissata.

Da segnalare che, normalmente, nelle stampanti seriali ad impatto la stampa avviene durante lo spostamento della testina da sinistra verso destra: questo significa che c'è un *tempo morto*, cioè in cui non viene stampato niente, che serve alla testina per riposizionarsi a capo. Naturalmente, questo pregiudica la velocità della stampante. In ogni caso ci sono oggi le cosiddette **stampanti bidirezionali**, che sono in grado di stampare indifferentemente da sinistra a destra ma anche da destra a sinistra, eliminando così il tempo morto e velocizzando la stampa.

Stampanti parallele ad impatto

Principale caratteristica di queste stampanti è quella di effettuare quasi contemporaneamente la stampa di una intera riga. I dispositivi più diffusi che

realizzano questo tipo di stampa sono il dispositivo "a rullo" e quello "a catena" o "a nastro".

Nelle **stampanti a rullo**, c'è un rullo cilindrico di acciaio sulle cui direttrici sono disposti i caratteri dell'alfabeto; il cilindro ruota con velocità costante attorno al proprio asse; ciascun carattere è ripetuto sulla generatrice del rullo tante volte quante sono le possibili posizioni di stampa e ci sono appositi martelletti che, battendo sul rullo, determinano la stampa dei caratteri nelle posizioni desiderate. Il numero massimo di caratteri per ogni riga varia fino ad un massimo di 160 caratteri; i valori più comuni sono comunque 120 e 132. Generalmente l'insieme dei caratteri di cui è dotata una stampante è costituito da 64 simboli grafici, 26 dei quali alfabetici, 10 numerici e i restanti simboli di interpunzione, di operazione e così via. Le velocità di funzionamento delle stampanti parallele ad impatto in genere va da 600 a 3000 righe al minuto.

Stampanti non ad impatto

I limiti delle stampanti ad impatto e le loro caratteristiche di costo, rumorosità e affidabilità hanno portato alla costruzione di **stampanti non ad impatto**: queste si basano su vari tipi di stampa che ora andremo a descrivere sinteticamente.

Un primo metodo è la **stampa termica**, che richiede l'uso di una stampa sensibile al calore sulla quale agiscono testine riscaldate elettricamente. E' un tipo di stampa poco diffuso, sia per le basse prestazioni sia soprattutto per la necessità di disporre di carta particolare. Tuttavia, è un sistema in generale poco costoso.

C'è anche la **stampa per elettroerosione**: la carta è ricoperta da una sottile pellicola di alluminio che viene fatta evaporare per punti mediante una scarica elettrica, scoprendo così la parte sottostante scura in modo da formare i caratteri voluti. Anche in questo caso, questo metodo si è poco diffuso, nonostante il prezzo relativamente limitato, a causa della carta particolare.

Notevole successo hanno avuto invece le **stampanti a getto d'inchiostro**, molto silenziose e potenti: in questo caso, un *ugello* emette, ad intervalli costanti di tempo, delle piccole *gocce* di inchiostro, alle quali viene attribuita una carica elettrostatica; le gocce vengono poi deflesse, lungo il loro tragitto, verso la carta da un campo elettrico e raggiungono la carta formando il carattere o il disegno voluto. Questo tipo di stampa può essere anche a colori.

I tre tipi di stampa non ad impatto appena descritti trovano la loro applicazione prevalentemente in *stampanti seriali* (lettera per lettera). C'è invece un altro tipo di stampa, la **stampa xerografica**, più spesso applicata alle *stampanti a pagina* con prestazioni elevate sia in termini di velocità sia in termini di qualità di stampa. Il principio è sostanzialmente lo stesso delle *fotocopiatrici*:

- per prima cosa, viene formata una *immagine elettronica*, per punti, della pagina da stampare;
- questa viene proiettata, mediante un raggio laser, su di un cilindro rivestito di un materiale fotoconduttore, che in precedenza è stato caricato elettrostaticamente;
- la carica presente sulla superficie del cilindro viene perciò neutralizzata "selettivamente" dalla luce e sul cilindro si forma così una "immagine elettrostatica" della pagina da stampare;
- questa "immagine elettrostatica" (dove cioè la pagina da stampare è costituita dai punti ancora carichi) viene tramutata in immagine effettiva mediante il deposito di uno strato di inchiostro (detto **toner**), anch'esso caricato elettrostaticamente, che aderisce al cilindro nei punti in cui la carica non era stata neutralizzata;
- infine, l'immagine costruita sul cilindro viene trasferita sulla carta e fissata su questa con un processo di fusione a seguito di riscaldamento.

La velocità di queste stampanti è dell'ordine di diverse centinaia di pagine al minuto; l'unico inconveniente è l'impossibilità di effettuare copie mediante carta carbone.

Registratori grafici o "plotter"

Spesso, può essere necessario costruire tramite un calcolatore dei disegni di buona qualità ed eventualmente anche di grandi dimensioni. Una unità di input adatta a questi scopi è il cosiddetto **registratore grafico** o **plotter**. Il *plotter* più diffuso è quello **a penna**, che costruisce il disegno desiderato mediante il semplice

movimento di una sola penna (o di più penne di colore diverso) su un supporto generalmente di carta.

MEMORIZZAZIONE DEI DATI

Generalità

Con il termine **memoria** si intende un dispositivo nel quale è possibile introdurre dei dati, conservarli ed estrarli dopo un certo tempo. *Memorizzare* un dato significa far assumere ai singoli componenti della memoria uno stato fisico che può essere messo in corrispondenza, mediante opportuni codici, con il dato stesso. L'elemento materiale nel quale vengono memorizzati i dati viene chiamato **supporto**. Una memoria può essere dunque vista come un insieme di elementi, generalmente chiamati **celle** (o *locazioni*), ciascuno dei quali è in grado di memorizzare un carattere o una parola di un dato *alfabeto*.

Per ragioni tecnologiche e di efficienza, è opportuno che le informazioni da memorizzare siano codificate mediante l'**alfabeto binario**: difatti, le memorie dei moderni calcolatori elettronici sono costituite da celle in grado di immagazzinare ciascuna una **cifra binaria** (o **bit**). Quindi, quando si parla di **capacità** di una memoria si intende il numero massimo di bit che possono essere immagazzinati nella memoria stessa.

Tuttavia, la capacità è espressa generalmente in unità multiple del bit: si tratta dei cosiddetti **byte**, che corrispondono ciascuno a 8 bit. Sappiamo che 8 bit sono sufficienti a memorizzare i caratteri di un alfabeto con 256 simboli distinti, oppure 2 cifre decimali

La lunghezza di una **parola** assume valori diversi (ad esempio 8 bit, oppure 16 oppure 32 o 36) nei vari calcolatori e essa coincide in genere con il numero di bit che vengono trattati contemporaneamente dall'**unità aritmetica e logica** del calcolatore. E' ovvio che il byte non è altro che una parola di 8 bit.

E' convenzione ormai accettata quella di esprimere la capacità totale della memoria in unità multiple del byte (o della parola), che sono il **Kilobyte** (1 kB = 1024 byte) e il **Megabyte** (1 MB = 1048576 byte). Dicendo quindi che la memoria di un calcolatore è ad esempio di 64K, si intende dire che essa contiene 65536 byte;

oppure, dicendo che la memoria è di 2 Mega, si intende dire che essa contiene 2097152 byte.

Un'altra importante caratteristica della memoria è la sua *velocità*, espressa generalmente in modo indiretto mediante il cosiddetto **tempo di accesso**: questo rappresenta l'intervallo di tempo che intercorre tra l'istante in cui il contenuto di una parola viene richiamato e quello in cui tale contenuto è disponibile. A questo proposito, è bene tenere presente che le memorie in uso attualmente possono essere classificate in 3 categorie:

- *memorie ad accesso uniforme*
- *memorie ad accesso diretto*
- *memorie ad accesso sequenziale*

Le **memorie ad accesso uniforme** sono quelle usate per la *memoria centrale* dei calcolatori e vengono così chiamate perchè ogni parola, a prescindere dalla lunghezza e dalla posizione occupata, ha lo stesso tempo di accesso. Il valore di questo tempo nei moderni calcolatori è dell'ordine di frazioni dei micro-secondi o addirittura dell'ordine dei nano-secondi (cioè i miliardesimi di secondo). Le memorie ad accesso uniforme sono anche dette **memorie RAM** (*Random Access Memory*), proprio per indicare che è possibile accedere ad ogni parola in memoria in un tempo che non dipende dalla parola stessa.

Le **memorie ad accesso diretto** sono basate sul movimento, in genere rotatorio, di alcuni elementi meccanici: il tempo di accesso dipende allora dalla posizione degli elementi nell'istante della richiesta di una informazione. Le memorie di questo tipo sono utilizzate come *memorie ausiliarie* di un calcolatore e il tempo di accesso va da un valore minimo dell'ordine del centesimo di secondo ad un valore massimo dell'ordine dei decimi di secondo (modelli peggiori).

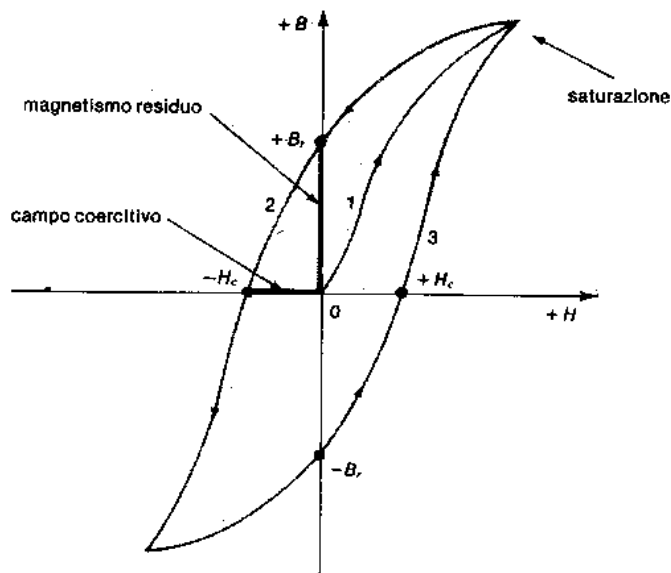
Le **memorie ad accesso sequenziale**, infine, sono basate su supporti che vengono esaminati sequenzialmente, uno dopo l'altro. In questo caso, non ha dunque senso parlare di tempo di accesso, in quanto non è logico accedere alle informazioni scegliendo un ordine sequenziale diverso da quello imposto dal supporto. Al contrario si può parlare di **velocità di lettura** (/scrittura) esprimendola in termini di bit letti (/scritti) al secondo.

Memorie a nuclei

Fino a qualche anno fa, la maggior parte dei calcolatori era dotato di memoria principale basata sull'impiego del cosiddetto **nucleo magnetico**, ossia un *anello di ferrite* avente il diametro di circa 1 mm attraversato da più fili elettrici. Attualmente, la tecnologia dei nuclei magnetici risulta superata, sia dal punto di vista economico sia da quello delle prestazioni: infatti quasi, tutti i calcolatori sono dotati di **memoria a semiconduttori**, dei quali parleremo più avanti. Nel seguito daremo comunque dei cenni sui principi di funzionamento delle **memorie a nuclei**.

Prima ancora di descrivere il modo in cui il nucleo magnetico può essere utilizzato come *cella di memoria*, ci sembra opportuno richiamare alcune considerazioni circa il comportamento del **materiale ferromagnetico** di cui è composto. Innanzitutto, cominciamo col dire che una sostanza si dice **ferromagnetica** quando, posta in un campo magnetico, è capace (come il ferro) di magnetizzarsi in modo particolarmente intenso e di conservarsi, almeno in parte, magnetizzata anche dopo che è cessata l'azione del campo magnetizzante.

Consideriamo il seguente diagramma cartesiano:



In ascisse mettiamo l'**intensità H di un campo magnetico**, mentre in ordinate mettiamo l'**induzione magnetica B** nel materiale ferromagnetico immerso nel campo.

L' *induzione magnetica* è in pratica il GRADO, o intensità, di magnetizzazione generato nel corpo soggetto al campo magnetico; essa cresce, non

proporzionalmente, al crescere dell'intensità H fino a raggiungere un certo valore, detto di **saturazione**, in corrispondenza del quale si ritiene totalmente realizzata la magnetizzazione del materiale (ossia la creazione definitiva di due polarità nel materiale considerato).

Vediamo allora più dettagliatamente cosa succede al variare dell'intensità H . Supponiamo innanzitutto che il materiale che stiamo considerando non sia mai stato magnetizzato. Partiamo dall'origine: siamo cioè in una situazione in cui sono nulle sia l'intensità del campo sia l'induzione magnetica nel materiale; adesso aumentiamo progressivamente l'intensità H del campo magnetico: l'induzione magnetica B cresce in modo abbastanza rapido fino ad arrivare ad un valore B_s , detto di **saturazione**, oltre il quale, pur aumentando ancora H , resta pressoché costante. In questo momento il materiale in questione è **magnetizzato**, ossia è dotato di due polarità.

Supponiamo adesso di fermarci al valore di saturazione di B al quale corrisponderà, sulle ascisse, un certo valore H_s del campo magnetico. Supponiamo quindi di ridurre gradatamente il valore del campo magnetico fino a riportarlo a zero; osserviamo anche qui cosa succede all'induzione B : essa, al diminuire di H , comincia sì a diminuire, ma non si annulla nel momento in cui $H=0$; al contrario, quando il campo magnetico risulta nullo, l'induzione assume un certo valore positivo $+B_r$, chiamato *magnetismo residuo* (o anche **induzione residua**). A questo punto, vogliamo fare in modo che l'induzione ritorni al valore 0: per farlo, dobbiamo abbassare ulteriormente il valore di H , dobbiamo cioè renderlo negativo: troviamo così che l'induzione B si annulla in corrispondenza di un valore (negativo) di H , che si indica con $-H_c$ e che prende il nome di **campo coercitivo**.

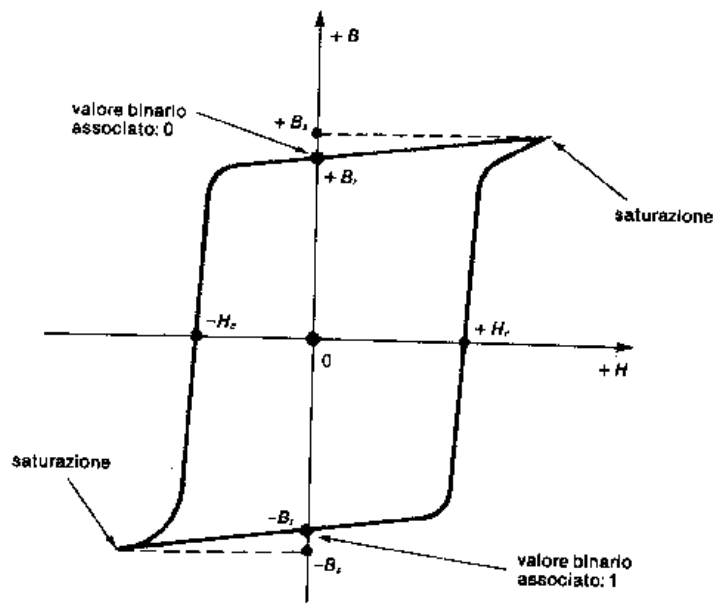
Ora proviamo a diminuire ancora H , cioè a renderlo sempre più negativo; osserviamo che anche l'induzione B diminuisce progressivamente (siamo nel semiasse negativo) fino ad un certo valore $-B_s$ di saturazione uguale ed opposto a quello raggiunto per valori positivi crescenti di H . Ci troviamo quindi in una situazione in cui sia H sia B hanno un valore negativo; riprendiamo adesso ad aumentare il valore H fino a renderlo di nuovo uguale a zero: si osserva che l'induzione magnetica B , in corrispondenza di $H=0$, non assume anch'essa valore 0, bensì assume un valore, negativo, $-B_r$, uguale ed opposto al valore del magnetismo residuo trovato per valori decrescenti positivi di H . Infine, se aumentiamo ulteriormente il valore del campo H , si verifica ovviamente che B cresce nuovamente fino allo stesso livello di saturazione trovato all'inizio del procedimento. Il valore del campo magnetico in corrispondenza del quale

l'induzione magnetica si annulla prende anche in questo caso il nome di *campo coercitivo*.

Il ciclo appena descritto prende il nome di **ciclo di isteresi** di un materiale ferromagnetico generico. I due valori di magnetismo residuo $+B_r$ e $-B_r$ (in corrispondenza dei quali c'è il campo H nullo) permettono quindi di stabilire il "senso" dell'ultimo campo magnetico applicato al materiale considerato; in modo analogo, i due valori $+H_c$ e $-H_c$, in corrispondenza dei quali si ottiene induzione B nulla, rappresentano le intensità ed i sensi dei campi magnetici coercitivi necessari per annullare i magnetismi residui.

Materiali particolarmente adatti per la costruzione di memorie sono le **ferriti** (ossidi di ferro), le quali sono infatti caratterizzate da valori delle induzioni magnetiche residue ben distinte e da altrettanto distinti valori dei campi coercitivi.

Un **nucleo di ferrite** ha un ciclo di isteresi la cui forma è praticamente rettangolare:

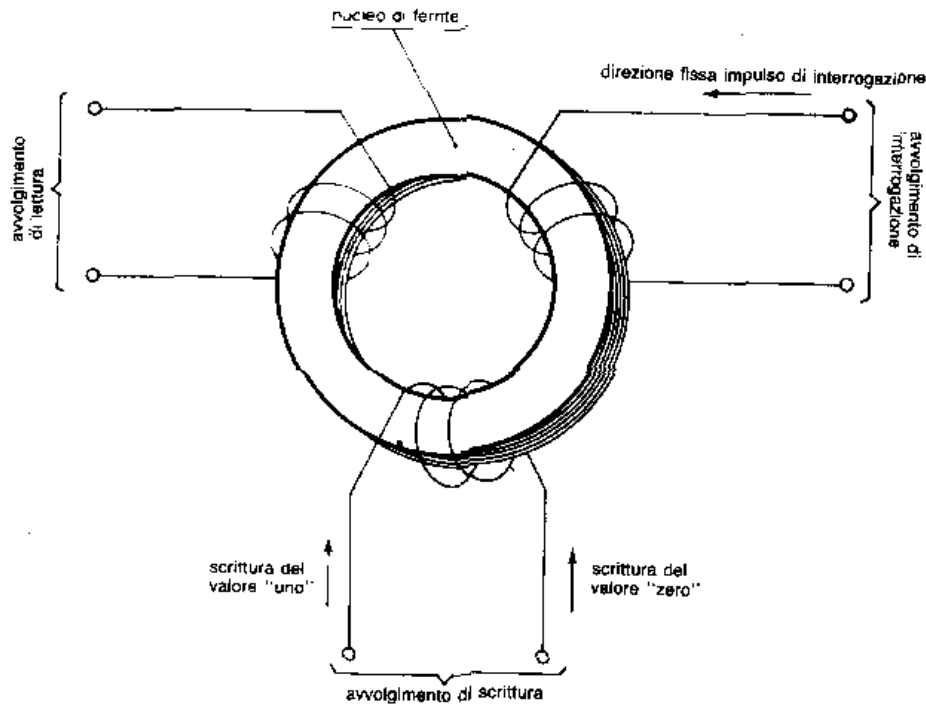


Ciclo di isteresi di un nucleo di ferrite

Osservando tale ciclo, si vede che il magnetismo residuo $+B_r$ è solo leggermente inferiore al valore massimo dell'induzione magnetica di saturazione $+B_s$. Questa importante caratteristica permette di realizzare nuclei magnetici aventi **due stati stabili di magnetizzazione** ben distinti, e cioè gli stati corrispondenti ai due valori $+B_r$ e $-B_r$ dell'induzione residua. E' allora evidente che elementi con tali caratteristiche di funzionamento sono tipicamente **binari** e perciò adatti ad essere

utilizzati come celle di memoria: basta stabilire una corrispondenza tra i due valori $+B_r$ e $-B_r$ e i due valori binari 0 e 1.

Per comprendere allora il principio di funzionamento di una memoria si consideri ora il seguente schema:



Schema di nucleo magnetico con avvolgimento

Abbiamo dunque un nucleo di ferrite (a forma di anello) attorno al quale si trovano tre **avvolgimenti** identici: tali avvolgimenti sono chiamati di **scrittura**, di **lettura** e di **interrogazione**.

Esistono due tipi di operazioni effettuabili su un nucleo: la *scrittura* di una informazione (ossia del valore binario 0 oppure 1) e la *lettura* per stabilire il valore dell'informazione in esso contenuta. Esaminiamo allora tali operazioni separatamente.

Scrittura di un nucleo magnetico

L'operazione di **scrittura** avviene inviando nell'avvolgimento di scrittura un impulso di corrente; tale impulso di corrente deve avere una intensità I_s tale da creare un campo magnetico di intensità H superiore al valore del campo coercitivo

H_c del nucleo. Allora, a seconda della **polarità** di tale corrente, il nucleo viene portato nel corrispondente stato di saturazione ($+B_s$ oppure $-B_s$) e questo avviene qualunque fosse il suo stato precedente (dato che il campo magnetico creato è superiore al campo coercitivo); in questo modo, alla fine dell'impulso, il nucleo, una volta cessato l'effetto del campo di intensità H (che scompare quando cessa l'impulso), avrà il valore di induzione magnetica residua ($+B_r$ oppure $-B_r$) cui corrisponde il valore binario che si intendeva scrivere (0 oppure 1)

Lettura di un nucleo magnetico

Fondamentalmente, la **lettura** di un nucleo magnetico si realizza inviando sull'avvolgimento di interrogazione un impulso di corrente (di polarità e intensità prestabilite) che, secondo la convenzione fatta, corrisponde alla scrittura di uno 0: si tratta perciò di un impulso di corrente in grado di generare un campo magnetico superiore al campo coercitivo H_c . Allora, ciò che interessa vedere, per capire quale valore c'è nel nucleo, è la *risposta* del nucleo stesso all'arrivo di questo impulso di corrente, in funzione del proprio stato di magnetizzazione. Possono perciò verificarsi due possibili casi:

- il nucleo si trova nello stato "0": in questo caso l'impulso di interrogazione non altera lo stato magnetico del nucleo; sull'avvolgimento di lettura vengono così generati due segnali, di piccola intensità, che sono dovuti alle variazioni di induzione magnetica dal valore $+B_r$ (corrispondente al valore binario 0) al valore di saturazione $+B_s$ e viceversa (osserviamo che i due segnali sono deboli in quanto abbiamo detto che questi due valori, $+B_r$ e $+B_s$, differiscono di poco);
- il nucleo si trova nello stato "1": adesso l'impulso di interrogazione altera lo stato magnetico del nucleo e lo porta al valore corrispondente allo "0". Sull'avvolgimento di lettura, in seguito alla variazione di flusso, viene quindi indotta una certa "tensione" il cui valore è proporzionale alla variazione del flusso.

In base, quindi, al valore della *tensione indotta*, sull'avvolgimento di lettura, dall'impulso di interrogazione, si deduce se il valore binario contenuto nel nucleo è 0 oppure 1. Se la tensione indotta è debole, allora si deduce che il nucleo contiene

il valore binario "0", mentre, se la tensione ha una certa intensità, allora si deduce che il nucleo conteneva il valore binario "1".

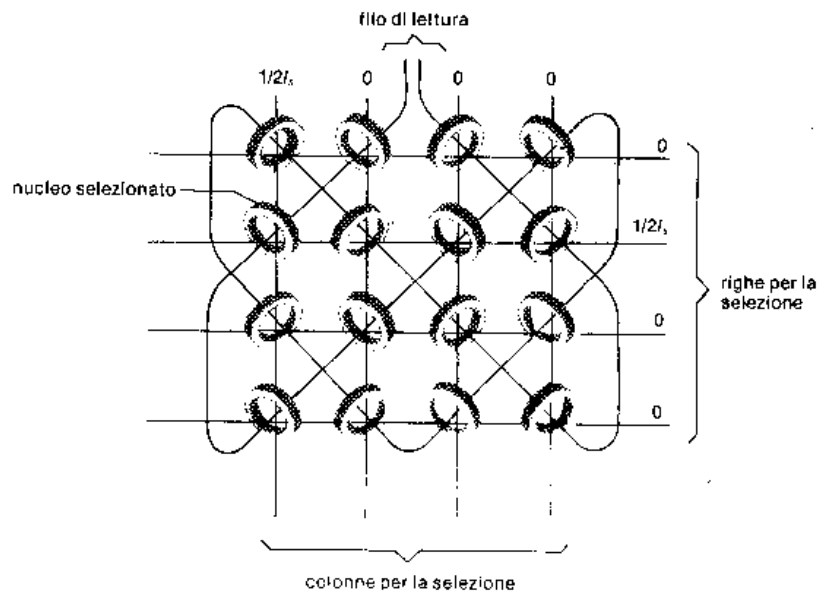
E' importante osservare che la lettura di un nucleo magnetico ha carattere **distruttivo**: infatti, l'informazione contenuta nel nucleo viene cancellata dal segnale di interrogazione inviato per la lettura. Perciò, alla fine di ogni lettura, il nucleo si trova sempre nello stato "0". Dato invece che spesso si vuole conservare una informazione pur dopo averla letta, dopo il ciclo di lettura il contenuto del nucleo viene opportunamente memorizzato e riscritto, nello stesso nucleo, mediante un normale ciclo di scrittura come quello analizzato in precedenza. Ecco perchè l'operazione di lettura di un nucleo prende il nome di **ciclo di lettura-scrittura**; il tempo impiegato per effettuare tale ciclo è detto **tempo di ciclo** della macchina.

Una caratteristica fondamentale delle memorie a nuclei magnetici è data dalla possibilità che esse hanno di conservare le informazioni memorizzate anche in caso di interruzione della tensione di alimentazione di tutto il calcolatore: tale proprietà si esprime dicendo che le memorie a nuclei non sono **volatili** ed è dovuta al fatto che le informazioni sono costituite dai segni delle induzioni magnetiche residue ($+B_r$ e $-B_r$) e queste sono indipendenti dalla presenza o meno della tensione. Se questo è un pregio delle memorie a nuclei, ci sono anche gli inevitabili inconvenienti: in primo luogo le loro dimensioni, poi l'elevato tempo di accesso ed i prezzi elevati dei sistemi per individuare (o meglio "selezionare") il nucleo sul quale si vuole operare.

Abbiamo poco fa usato il termine "selezionare": infatti, la memoria di un calcolatore è costituita da una infinità di nuclei ed è perciò necessario trovare un modo per considerarne uno solo posto in una posizione qualsiasi. Vediamo allora in che cosa consiste la **selezione** di un nucleo di memoria.

Il principio generale per la selezione del nucleo sul quale si vuole operare è quello detto **per coincidenza** e sfrutta la particolare forma del *ciclo di isteresi*. Consideriamo un conduttore che attraversa il nucleo considerato; se facciamo passare, attraverso tale conduttore, una corrente di intensità pari a metà di quella necessaria per la saturazione, questa corrente ha un effetto molto limitato sullo stato di magnetizzazione del nucleo: infatti, provoca solo un modesto segnale di disturbo sul filo di lettura. Allora, per realizzare la *commutazione* del nucleo, si prendono due fili (ad esempio quello di scrittura o quello di interrogazione) che passano attraverso il nucleo e si inviano tramite essi delle correnti di intensità pari a $\frac{1}{2}I_S$, dove con I_S avevamo indicato il valore necessario per portare l'induzione

magnetica B al livello di saturazione B_s . Supponiamo adesso che i nuclei siano disposti in modo da costituire un reticolo a due dimensioni:



A ciascun nucleo è possibile associare due numeri che indicano la riga e che la colonna alle quali esso appartiene. In questo modo la selezione di $m \cdot n$ nuclei (quanti cioè ce ne sono in una "matrice" di m righe per n colonne) è ricondotta alla selezione di $m+n$ conduttori, se due dei quali (passanti naturalmente per il nucleo prescelto) si inviano correnti di intensità $\frac{1}{2}I_s$. Mandando tali correnti si verifica la seguente situazione:

- solo nel nucleo selezionato si produrrà un campo magnetico di intensità uguale a quello prodotto dalla corrente I_s , in quanto esso viene attraversato da entrambi i conduttori sui quali è stato inviato un impulso pari a $\frac{1}{2}I_s$;
- in tutti gli altri nuclei della stessa riga e della stessa colonna, si produrrà un campo magnetico di intensità corrispondente a $\frac{1}{2}I_s$, in quanto è uno solo il conduttore che passa per tali nuclei, ed è perciò insufficiente per provocare variazioni della magnetizzazione residua.

Secondo le convenzioni stabilite in precedenza, la scrittura di uno "0" viene effettuata inviando sui due conduttori che individuano il nucleo prescelto delle correnti di intensità $\frac{1}{2}I_s$; la scrittura di un "1" avviene invece inviando correnti di

intensità pari a $-\frac{1}{2}I_s$; infine l'interrogazione avviene con lo stesso procedimento della scrittura.

Se osserviamo la disposizione a matrice dei nuclei magnetici, notiamo facilmente che il filo di lettura attraversa tutti i nuclei della matrice non a caso, bensì secondo un *tracciato* particolare: tale tracciato è stato concepito in modo da effettuare una "compensazione" dei segnali di disturbo che vengono generati, come abbiamo detto prima, dai nuclei interessati da correnti di intensità $\frac{1}{2}I_s$.

Il metodo appena esposto consente di "selezionare" un solo nucleo per volta e si dice perciò che opera con **grado di parallelismo pari a 1**; per aumentare la velocità di funzionamento, è possibile operare anche su più nuclei contemporaneamente, in modo da realizzare un grado di parallelismo maggiore di 1: per farlo i nuclei vengono disposti secondo un reticolo non più di 2 ma di 3 dimensioni: si ha cioè una serie di *reticoli bidimensionali* giacenti su tanti piani paralleli quanto è il grado di parallelismo desiderato: se ad esempio si vuole un grado uguale a 5, ci sono 5 reticoli bidimensionali (cioè 5 matrici di nuclei) disposti parallelamente uno rispetto all'altro.

Memorie a semiconduttori

Diamo adesso dei brevissimi cenni circa le **memorie a semiconduttori**. La realizzazione di queste memorie è stata successiva a quella delle memorie a nuclei e la tecnologia richiesta è più sofisticata; tuttavia, queste memorie hanno ormai soppiantato del tutto le memorie a nuclei, salvo qualche rara eccezione (in particolare le memorie a nuclei, per il fatto di conservare le informazioni anche in mancanza di tensione, sono ancora utilizzate nelle navicelle spaziali).

Le memorie a semiconduttori presentano notevoli vantaggi: in primo luogo, un elevato *livello di integrazione* (fino ad 1 milione di bit per ogni "chip"); inoltre, presentano un basso costo unitario e una elevata velocità di lettura e scrittura; infine, la possibilità di rendere molto più semplice la realizzazione dei circuiti che effettuano l'indirizzamento e la selezione.

Memorie ausiliarie o di massa

L'elevato costo di produzione delle memorie descritte in precedenza rende poco conveniente il loro impiego per contenere tutte le informazioni che sono necessarie per la risoluzione, mediante calcolatore, di problemi complessi. Per questo, nella maggior parte dei calcolatori vengono utilizzate, oltre alla *memoria principale* (che è ad accesso uniforme) avente capacità limitata ma elevatissima velocità di funzionamento, altri tipi di memorie, dette **memorie ausiliarie** o anche **memorie di massa**, che hanno prestazioni meno elevate ma, in compenso, una più elevata capacità di memorizzazione.

Il sistema generalmente utilizzato in questi tipi di memorie prevede che le singole informazioni siano memorizzate magnetizzando delle areole su superfici ricoperte da un sottile strato di materiale ferromagnetico (del ferromagnetismo abbiamo già parlato a proposito delle memorie a nuclei di ferrite). Queste areole non sono altro che le celle delle memorie.

Le memorie di massa che funzionano con questo metodo di magnetizzazione hanno tutte una caratteristica fondamentale che riguarda l'accesso alle varie celle: è il supporto della memoria stessa che effettua un movimento meccanico rispetto al dispositivo che si occupa della lettura e della scrittura. Tale movimento provoca il passaggio delle celle davanti al *dispositivo di lettura-scrittura* e quindi, ad esempio nella lettura, la generazione di un segnale che permette di individuare l'informazione contenuta nella cella esaminata.

Il *dispositivo di lettura-scrittura* non è altro che una **testina magnetica** sul cui *avvolgimento* vengono inviati i segnali per la scrittura e vengono indotti i segnali per la lettura delle informazioni registrate in precedenza.

E' importante notare che la lettura di questo tipo di memorie non è di tipo *distruttivo*, ossia non altera in nessun modo il contenuto delle celle esaminate.

Esistono infine vari tipi di memorie di massa nelle quali il meccanismo di lettura-scrittura è del tipo appena descritto; esse variano a seconda del meccanismo di movimento, della velocità di funzionamento e del modo di accesso.

Nastri magnetici

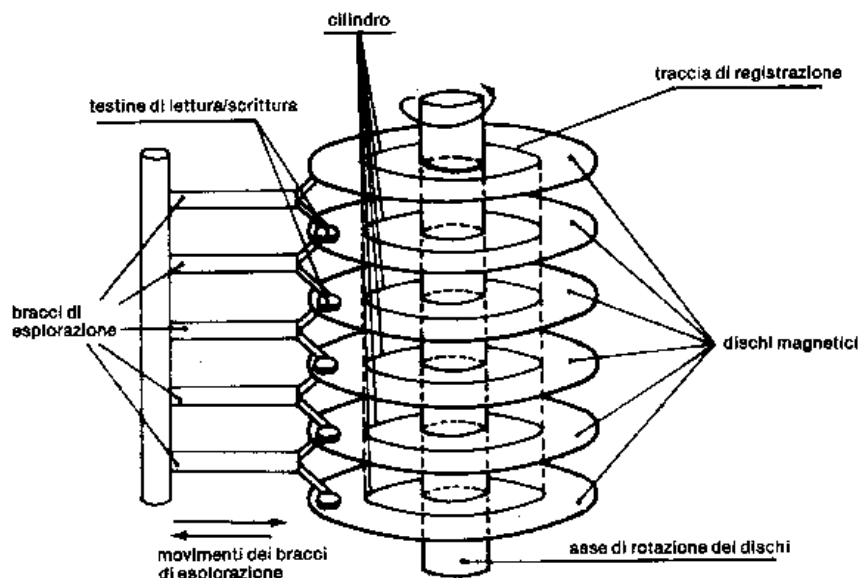
Il **nastro magnetico** è una memoria di massa ad *accesso sequenziale*; esso è costituito da un supporto di materia plastica, di larghezza fissa (mezzo pollice) e lunghezza variabile, avente una faccia ricoperta da un sottile strato uniforme di

ossido di ferro, che è una sostanza ferromagnetica. La superficie di un nastro magnetico può essere considerata come suddivisa in tante piste longitudinali, chiamate **canali** o anche **tracce**, di lunghezza uguale a quella del nastro e di larghezza sufficiente per contenere una **cella**; ogni cella, come sappiamo, può essere poi magnetizzata in due stati opposti cui sono associati i valori binari "0" e "1". Il nastro magnetico viene raccolto su bobine montate su un apposito dispositivo, detto **unità nastro**, che ne comanda l'avvolgimento e lo svolgimento, permettendo così il passaggio sotto le **testine di lettura e di scrittura**.

La **densità di registrazione** (o *impaccamento*) si misura in termini di numero di bit per pollice memorizzati in una sola traccia (brevemente in **bpi**). La velocità di lettura (quantità di caratteri che possono essere scritti o letti in un secondo) dipende sia dall'impaccamento sia dalla velocità di trascinamento del nastro.

Dischi magnetici

Anche le **memorie a dischi magnetici** appartengono alla categoria delle memorie di massa ad accesso diretto. Esse sono costituite da un insieme di dischi coassiali, costituenti una "pila", opportunamente distanziati l'uno dall'altro e ruotanti solidariamente con velocità uniforme attorno all'asse comune:



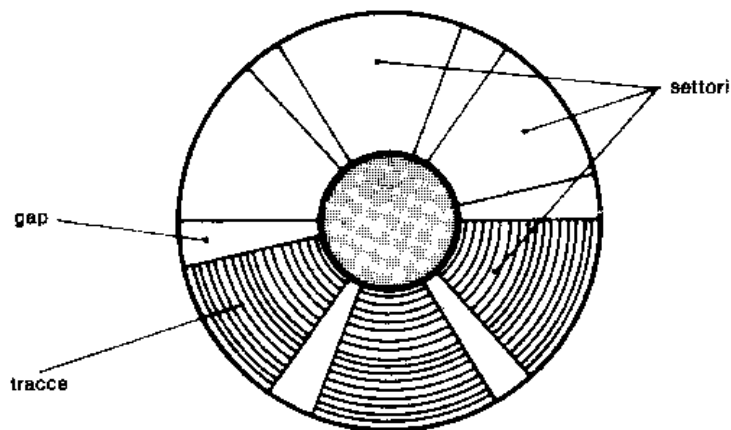
Schema di una pila di dischi magnetici

Le facciate di ogni disco sono ricoperte da uno strato di **materiale ferromagnetico**; in corrispondenza di ogni facciata, si trova una **testina unica di lettura-scrittura**, capace di movimento radiale o verso l'interno o verso l'esterno del disco. Le testine si trovano su altrettanti *bracci meccanici*, detti di **esplorazione**, che si muovono solidariamente come un "pettine porta testine". Le facciate esterne di una pila di dischi, cioè la facciata superiore e quella inferiore, non vengono solitamente utilizzate per la conservazione dei dati e questo per motivi di sicurezza dei dati.

L'apparecchio che si occupa del movimento delle pile di dischi e del funzionamento delle testine prende il nome di **unità dischi**.

Per quanto riguarda i principi fisici utilizzati per la registrazione e le lettura delle informazioni sono gli stessi usati nei nastri e nei tamburi magnetici.

Ogni faccia di ogni disco è divisa in un numero variabile (a secondo del modello) di **tracce** concentriche: si va da 800 tracce per i modelli più piccoli a 1200 per quelli più capienti. Le tracce sono a loro volta divise in **blocchi**, separati da zone vuote dette **gap**. I blocchi possono essere sia di lunghezza variabile sia di lunghezza fissa: in quest'ultimo caso si chiamano **settori** e i gap si trovano in posizione fisse sulle singole tracce. Chiaramente, il contenuto di un blocco può essere letto o riscritto solo se il blocco stesso passa sotto la testina di lettura-scrittura.



Il **tempo di accesso ad un blocco** è variabile ed è il risultato della somma dei seguenti tempi:

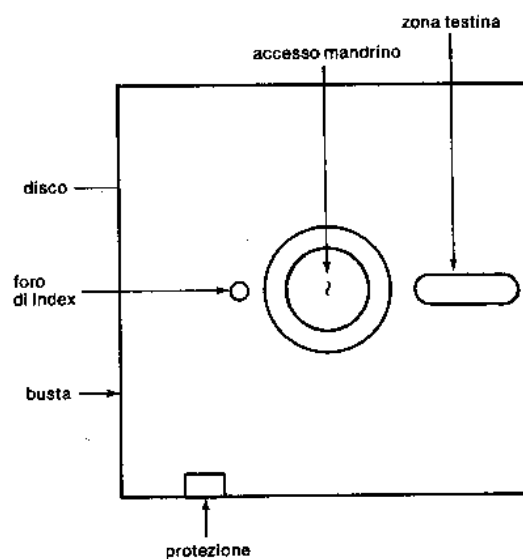
- **tempo di posizionamento** della testina sulla traccia contenente il blocco interessato; questo tempo dipende sia dal modello di unità a disco sia dalla distanza che la testina deve percorrere per trovare il blocco interessato;
- **tempo di latenza** necessario per ottenere il passaggio del blocco cercato sotto la testina; in media questo tempo è pari al semiperiodo di rotazione dei dischi;
- **tempo di lettura** vero e proprio.

La **capacità** di una pila di dischi può variare da qualche milione di caratteri fino a diversi miliardi.

Supporti magnetici di basso costo

La diffusione dei PC ha favorito la diffusione di dischi e nastri magnetici di costo e prestazioni modeste che hanno trovato applicazione anche come supporti di input-output. Le apparecchiature per la registrazione e la lettura di dati su questi supporti magnetici sono concettualmente simili a quelle già viste, tuttavia i *dispositivi di trascinamento* sono molto più semplici e ovviamente permettono il raggiungimento di prestazioni molto più limitate.

Un supporto economico analogo ai dischi magnetici è il **floppy disk**: si tratta di un disco flessibile che, a seconda dei modelli, può avere un diametro compreso tra circa 8 e 20 centimetri.



Il disco flessibile è di plastica rivestita di *materiale ferromagnetico* ed è contenuto permanentemente in una busta quadrata sempre di plastica; c'è un foro centrale per il dispositivo di trascinamento ed altre aperture che permettono alle testine di effettuare le operazioni di lettura e di registrazione.

Le tracce sono in generale suddivise in *settori* di lunghezza fissa di 256 o 512 byte e la capacità totale varia da 250.000 a un milione di byte.

Intermedi tra i floppy disk e i dischi magnetici di grande capacità visti in precedenza sono i cosiddetti **hard disk**, che hanno dimensioni analoghe ai floppy; essi, essendo racchiusi insieme con le testine di lettura in un contenitore ermetico, offrono prestazioni molto migliori, sia in termini di capacità di memorizzazione sia in termini di velocità di accesso.

Altri tipi di memorie

Data l'importanza fondamentale che la memoria riveste per un calcolatore elettronico, ci si sforza continuamente di trovare delle soluzioni che permettano alte prestazioni per quanto riguarda la capacità di memorizzazione e la velocità di accesso. Attualmente, il filone più promettente nell'area delle memorie di massa sembra essere quello delle **memorie ottiche**, basate sulla modifica permanente di un supporto provocata da un *raggio laser* e sulla possibilità di riconoscere che tale modificazione è avvenuta, sempre per mezzo di un fascio di luce. I dischi usati in questo caso sono i **Compact Disk**, i quali forniscono prestazioni molto interessanti accompagnate però da difetti non irrilevanti, almeno allo stato attuale di tecnologia.

Per quanto riguarda le capacità di memorizzazione, un solo disco del diametro di 12 cm può immagazzinare fino a 500 milioni di caratteri, offrendo al tempo stesso una notevole sicurezza dei dati. La limitazione di questi dispositivi consiste, almeno per il momento, nella impossibilità di modificare o cancellare le informazioni contenute, per cui tali dischi si possono usare come supporti di memoria a sola lettura (da cui il termine **CD-ROM**).

FUNZIONI DI CALCOLO E DI GOVERNO

Le istruzioni

Le *funzioni di calcolo e di governo*, in un calcolatore elettronico, vengono svolte dall'unità centrale del calcolatore, detta **Central Processing Unit (CPU)**: essa comprende l'*unità aritmetica e logica*, l'*unità di governo vera e propria* e la *memoria principale*. L'unità di governo e quella aritmetica e logica sono in grado di accedere soltanto ai dati che sono contenuti nella memoria principale: questo comporta che qualsiasi dato residente in *memorie ausiliarie* deve necessariamente essere trasferito nella memoria centrale per poter essere elaborato. Oltre a questo, anche la trasmissione dei dati dalle unità di input alle memorie ausiliarie e, viceversa, dalle memorie ausiliarie alle unità di input deve avvenire tramite la memoria centrale.

La memoria centrale può contenere sia i **dati** (già elaborati o da elaborare) sia la descrizione dell'*algoritmo* da eseguire. Tale descrizione viene fornita mediante una successione di *ordini elementari* codificati in modo che l'unità di governo possa correttamente interpretarli: si tratta delle cosiddette **istruzioni** e, più precisamente, delle **istruzioni macchina**, per distinguerle dalle **istruzioni simboliche** di cui si parlerà più avanti.

L'insieme delle istruzioni che realizzano un determinato algoritmo costituiscono un **programma**. I programmi composti da istruzioni macchina si dice che sono stati redatti in **linguaggio macchina**, per distinguerli dai programmi composti da *istruzioni simboliche*.

Le istruzioni macchina che compongono un programma vengono registrate nella memoria centrale in modo sequenziale; esse occupano, nella maggior parte dei casi, uno spazio notevole della memoria; in ogni caso, si tratta di una occupazione temporanea in quanto il calcolatore è comunque in grado di ricevere nuovi algoritmi e di sostituirli a quelli già presenti.

Una istruzione macchina si considera divisa in due parti fondamentali:

- la prima di queste serve a caratterizzare il tipo di operazione che deve essere eseguita e prende il nome di **codice operativo**;

- la seconda invece individua la posizione dei dati, su cui tale operazione deve essere eseguita, nella memoria centrale: a tali dati si dà il nome di **operandi**.

Per individuare gli *operandi* è necessario associare ad ogni cella della memoria (o ad ogni parola) un codice che permetta al calcolatore di individuarla in modo univoco. Questo codice si chiama **indirizzo**. La corrispondenza biunivoca tra indirizzo e cella di memoria si realizza, dal punto di vista pratico, con un *sistema di circuiti* che permette di registrare o leggere il contenuto di una cella che si trova ad un determinato indirizzo.

L'**indirizzamento della memoria centrale** si ottiene stabilendo una corrispondenza biunivoca tra l'insieme delle celle che costituiscono la memoria stessa e l'insieme dei numeri naturali minori del numero che esprime la capacità della memoria (incluso lo zero). Facciamo un esempio: supponiamo che il calcolatore preso in considerazione abbia una memoria a nuclei magnetici costituita da 4096 byte: allora, ognuno di questi byte avrà un numero naturale (compreso tra 0 e 4095) di riferimento, in modo che a ciascun numero corrisponda un solo byte e viceversa. Da notare che il numero 4096 non è casuale: infatti, dato che motivi tecnologici impongono l'utilizzo dell'alfabeto binario per un calcolatore, tutta l'organizzazione del calcolatore si fonda sull'alfabeto binario, per cui la stessa capacità della memoria viene espressa da un numero che sia una potenza di 2 ed infatti $4096 = 2^{12}$.

Torniamo adesso alle istruzioni macchina. Ciascuno degli operandi consiste essenzialmente di un *indirizzo* (quello dei dati su cui si deve eseguire l'algoritmo) e, eventualmente, è completato dall'indicazione della **lunghezza** dell'operando, nel caso che questo possa occupare un numero variabile di byte (o parole).

Nei casi concreti, gli indirizzi specifici (o, ciò che è lo stesso, gli operandi) non sono mai più di tre, in quanto si può verificare che con 3 operandi è possibile specificare ogni cosa riguardo una qualsiasi operazione; talvolta gli operandi possono anche essere 2 o anche 1 solo o addirittura l'istruzione potrebbe essere priva di operandi. Ad esempio, una istruzione di sottrazione, che dovrebbe intuitivamente constare di 3 operandi (1 per il minuendo, 1 per il segno di sottrazione e 1 per il sottraendo) potrebbe invece avere due soli operandi: uno per il minuendo con incorporato il segno di sottrazione e l'altro per il sottraendo; potrebbe anche avere uno solo operando, per il sottraendo: in questo caso, si conviene che sia il segno sia il minuendo si trovino in un dispositivo, detto

accumulatore, il cui indirizzo è implicito; questo significa che, una volta dato il codice operativo e il valore del sottraendo, il calcolatore sa immediatamente dove deve andare a prendere il minuendo e che deve fare la sottrazione.

E' il caso di osservare che in molti calcolatori esistono in realtà una serie di accumulatori, che prendono il nome di **registri**: in questo caso, al posto di specificare un indirizzo di memoria, nella istruzione va specificato il numero del registro in cui si trova l'operando da considerare.

Di norma, l'**interpretazione** e l'**esecuzione** delle istruzioni che compongono un programma avvengono nell'ordine in cui le istruzioni stesse sono registrate in memoria e a partire da un indirizzo che viene fissato di volta in volta (che cioè cambia da programma a programma). Tuttavia esistono delle particolari istruzioni, dette di **salto**, che permettono di alterare l'ordine di esecuzione, per esempio interrompendo la sequenza in corso per iniziarne un'altra registrata in una qualsiasi posizione nella memoria centrale.

Il **repertorio** delle istruzioni di un calcolatore, ossia l'insieme di tutte le istruzioni (con diverso codice operativo) che il calcolatore può eseguire, varia da un calcolatore all'altro; in ogni caso, si possono individuare le seguenti categorie di istruzioni che sono sempre presenti:

- *istruzioni aritmetiche e di confronto*
- *istruzioni di input e di output*
- *istruzioni di controllo (salto)*

Alla prima categoria si possono anche aggiungere le cosiddette *istruzioni di trasferimento*, le quali trasportano dati da una zona della memoria principale ad un'altra.

Le **istruzioni di confronto** permettono di stabilire, dati due operandi A e B, quale delle seguenti relazioni è verificata:

$$A > B \quad A < B \quad A = B$$

Una volta effettuato il confronto, il risultato viene memorizzato, sempre tramite opportuni codici, in particolari posizioni della memoria centrale o anche in memorie specializzate (in modo da distinguerlo da tutti gli altri dati presenti).

Le **istruzioni di I/O** permettono il trasferimento dei dati da una unità di input alla memoria principale o, viceversa, dalla memoria centrale ad una unità di output.

Infine, le **istruzioni di controllo** sono essenzialmente comandi di salto che, come già detto, permettono di alterare l'ordine con cui devono essere eseguite le istruzioni che compongono il programma: questa alterazione dell'ordine può avvenire incondizionatamente oppure solo se si verificano certe condizioni, per esempio se si ottiene un particolare risultato o se risulta vera una particolare proposizione. Naturalmente, le istruzioni di salto contengono l'indirizzo della posizione di memoria dell'istruzione che si vuole eseguire appena terminato il salto.

Per concludere, un'ultima osservazione: ad ogni istruzione, come a qualsiasi dato registrato nella memoria centrale, viene associato un indirizzo che permetta ad altre istruzioni di fare riferimento ad essa. Bisogna però sottolineare che gli elementi costitutivi di una istruzione, e cioè il codice operativo e i vari operandi, sono indistinguibili dai dati comuni nel momento in cui vengono registrati nella memoria centrale: ecco perchè è sempre importante specificare a che punto della memoria centrale il calcolatore deve cominciare a leggere e interpretare le istruzioni; non specificandolo, il calcolatore andrebbe a leggere i dati e cercherebbe, invano, di interpretarli come istruzioni.

L'unità di governo

L'**esecuzione** di una *istruzione in linguaggio macchina* comporta l'esecuzione di un numero (variabile) di operazioni; tali operazioni sono descritte mediante dei comandi elementari che vengono interpretati dai circuiti che compongono l'**unità di governo** del calcolatore. Nei primi calcolatori, ad ogni istruzione veniva associata una *rete logica di circuiti* in grado di svolgere le operazioni richieste. Più tardi, viste la complessità e l'abbondanza delle istruzioni eseguibili da un calcolatore, si è giunti al concetto della cosiddetta **microprogrammazione**: in breve, la rete logica che serve ad eseguire le operazioni è stata parzialmente sostituita da una *unità di memoria*, nella quale sono state memorizzate le sequenze di comandi elementari chiamate **microprogrammi**; quando c'è da eseguire una certa istruzione in linguaggio macchina, ad essa vengono associati gli indirizzi dei vari microprogrammi che servono per effettuare le operazioni da essa richieste. I comandi elementari che compongono i microprogrammi vengono detti

microistruzioni e la memoria utilizzata per conservarli viene chiamata **memoria di controllo**.

L'insieme delle microistruzioni di cui è dotato un calcolatore viene comunemente indicato con il termine **firmware** e costituisce l'interfaccia tra il *software* (cioè l'insieme dei programmi di cui dispone il calcolatore) e l'*hardware* (cioè l'insieme di apparecchiature).

La funzione principale dell'unità di governo è quindi quella di generare, per ogni istruzione macchina, delle sequenze di comandi elementari la cui esecuzione realizza le operazioni richieste. Osserviamo però che l'esecuzione dei comandi elementari non è un processo continuo, bensì avviene per "passi" o **cicli di macchina**. La durata di ogni ciclo viene determinata dal **clock** interno al sistema.

L'istruzione originale viene in primo luogo codificata; poi viene interpretata e quindi, in base al suo codice operativo, viene determinato un indirizzo di inizio del microprogramma corrispondente. Oltre ai circuiti per la codifica e la decodifica degli indirizzi corrispondenti all'istruzione macchina considerata, l'unità di governo è costituita quindi essenzialmente da una *memoria di controllo* nella quale vengono memorizzati i microprogrammi: si tratta di una memoria "a sola lettura" (**Read Only Memory**), il che garantisce la massima sicurezza delle informazioni contenute, ma anche la possibilità di ottenere alte prestazioni a costi contenuti.

Funzionamento dell'unità di governo

Vediamo adesso, a grandi linee, il funzionamento dell'unità di governo:

- 1) per prima cosa, l'unità di governo preleva, dalla memoria principale, l'istruzione che deve essere eseguita. L'indirizzo di memoria nel quale si trova tale istruzione può trovarsi in due possibili "siti": potrebbe trovarsi in un **Registro Contatore delle istruzioni** (detto anche **Program Counter**), oppure potrebbe trovarsi, insieme ad altri dati sulla situazione del programma in corso di esecuzione, in un **Registro Stato del Programma**, del quale parleremo tra poco;
- 2) dopo averla prelevata, l'unità di governo deposita l'istruzione da eseguire in un proprio **Registro Istruzione Corrente**;
- 3) a questo punto, l'istruzione da eseguire viene decodificata e, attraverso il suo codice operativo, viene subito individuato l'indirizzo della prima

microistruzione da eseguire (indirizzo che è depositato nell'apposito registro);

- 4) nel momento in cui l'orologio del sistema abilita l'uscita di questo registro, si ha l'accesso alla memoria ROM; da questa si ottengono in uscita sia i comandi corrispondenti alla microistruzione da eseguire sia l'indirizzo della prossima microistruzione (che verrà memorizzato nell'apposito registro).

Questa sequenza di operazioni si ripeterà fino alla fine del microprogramma, al termine del quale l'unità di governo passa a considerare l'istruzione successiva.

Possiamo a questo punto suddividere l'esecuzione di una istruzione in due fasi fondamentali: la prima è la fase di **interpretazione** e la seconda è la fase della vera e propria *esecuzione*. Nella prima fase sono comprese tutte le operazioni, appena descritte, svolte dall'unità di governo; la seconda fase viene effettuata dall'unità aritmetica e logica oppure dalle unità di controllo delle periferiche.

Per le operazioni di interpretazione e di esecuzione di un programma, viene normalmente utilizzato il già citato **Registro Stato del Programma (RSP)**, detto anche **Program Status Word (PSW)**; esso contiene un certo numero di informazioni essenziali sulla situazione del programma che è in fase di esecuzione; in particolare contiene:

- il codice operativo dell'istruzione che è in corso di esecuzione;
- l'indirizzo della istruzione che dovrà essere eseguita al termine dell'istruzione in corso;
- la lunghezza dell'istruzione in corso;
- le condizioni poste dall'ultima istruzione di confronto o istruzione aritmetica eseguita;
- eventuali segnalazioni di errore.

Ci sono poi altre informazioni contenute nel RSP, come ad esempio quelle connesse con i *meccanismi di interruzione del programma*, ma ne parleremo in altra sede.

La consultazione e la gestione delle informazioni contenute nel RSP si svolgono secondo il seguente schema:

- 1) prima ancora di avviare l'esecuzione del programma registrato nella memoria principale, l'indirizzo della prima istruzione del programma (detto indirizzo di lancio) viene scritto in quella zona del RSP che contiene anche l'indirizzo dell'istruzione successiva;
- 2) l'istruzione il cui indirizzo è contenuto nel RSP viene prelevata dalla memoria e viene inserita nel Registro Istruzione Corrente (il tutto, come abbiamo visto, a cura dell'unità di governo);
- 3) viene quindi avviata la fase di interpretazione dell'istruzione;
- 4) se l'istruzione viene interpretata nel modo corretto, la sua lunghezza (che dipende esclusivamente dal codice operativo) viene scritta nell'apposita zona del RSP; poi essa viene sommata all'indirizzo registrato nello stesso RSP; si ottiene così un nuovo indirizzo che risulta essere quello della istruzione da eseguire successivamente, a meno che l'istruzione in corso non sia una istruzione di salto;
- 5) viene quindi avviata la fase di esecuzione dell'istruzione in corso; se si tratta di una istruzione di salto e, inoltre, se la condizione cui è eventualmente condizionata risulta verificata, l'indirizzo dell'istruzione successiva contenuto nel RSP viene sostituito con l'indirizzo indicato dalla istruzione stessa;
- 6) completata la fase di esecuzione, il calcolatore è in grado di iniziare l'interpretazione della istruzione successiva in quanto l'indirizzo di questa si trova già nel RSP.

Il ciclo di operazioni descritto ai punti 2 e 6 si può adesso ripetere in modo del tutto automatico e a velocità elevatissime. L'**arresto** del programma può avvenire quando viene eseguita una istruzione apposita oppure quando una determinata istruzione, per un motivo o per un altro, non può essere interpretata o eseguita in modo corretto. In quest'ultimo caso, si dice che il calcolatore si trova in una situazione di **program check**, ossia una situazione in cui i dispositivi e le procedure di controllo hanno rilevato e segnalato una situazione anomala. L'arresto indesiderato di un programma potrebbe essere ad esempio dovuto al fatto che una istruzione si trova a dover utilizzare indirizzi non ammessi (ad esempio maggiori della capacità della memoria).

L'unità aritmetica e logica

Dal punto di vista strettamente funzionale, l'**unità aritmetica e logica** va considerata distinta dall'*unità di governo*; tuttavia gli sviluppi tecnologici attuali tendono invece a integrare queste due unità, che appartengono entrambe all'unità centrale del calcolatore (CPU).

La funzione specifica dell'unità aritmetica e logica è quella di eseguire le operazioni aritmetiche, logiche e di confronto sui dati che si trovano registrati nella memoria principale.

Di norma, un calcolatore dispone di apposite istruzioni per l'esecuzione delle quattro operazioni aritmetiche fondamentali sia sui dati decimali sia su quelli di tipo binario; tuttavia, calcolatori di piccole dimensioni possono essere dotati delle sole istruzioni per l'addizione e la sottrazione. In questi casi particolari quanto rari, la moltiplicazione e la divisione vengono realizzate mediante sequenze di istruzioni (*sottoprogrammi*), registrate nella memoria principale, che eseguono rispettivamente una serie di addizioni e sottrazioni. Questo metodo viene anche utilizzato, mediante gli opportuni algoritmi, per la realizzazione di operazioni più complesse, come ad esempio l'elevamento a potenza o l'estrazione di radice quadrata o il calcolo delle funzioni trascendenti più comuni: per poter aumentare considerevolmente la velocità di esecuzione, molte di queste funzioni vengono realizzate mediante opportuni **microprogrammi** che contribuiscono ad aumentare il firmware del sistema.

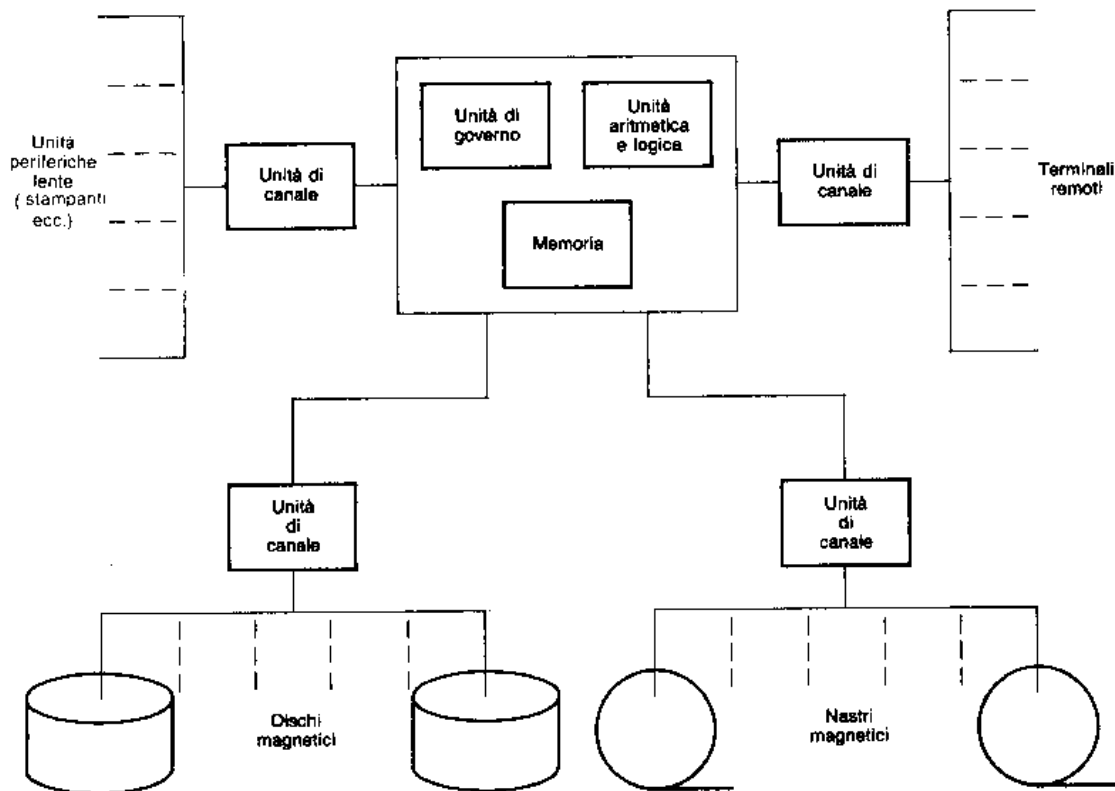
Per finire, indichiamo le principali componenti di una unità aritmetica e logica:

- la **rete logica di calcolo**, che è costituita da vari circuiti elettronici elementari che eseguono le operazioni logiche fondamentali (AND, OR, NOT), le operazioni aritmetiche semplici e i confronti sui dati contenuti nella memoria principale;
- uno o più registri, chiamati **accumulatori**, utilizzati per contenere uno o più dati a seconda del tipo e del formato dell'istruzione che è in corso di esecuzione.

Cenni sulle architetture dei sistemi di elaborazione

Non è possibile in questa sede sviluppare in modo adeguato l'argomento delle **architetture**, per cui esporremo solo concetti fondamentali.

Un primo tipo di architettura, spesso adottato nei calcolatori medi e grandi, comprende una serie di **unità di canale** che provvedono a realizzare i collegamenti tra l'unità centrale e le varie unità periferiche (quali stampanti, terminali remoti, nastri magnetici o dischi magnetici, cioè memorie ausiliarie):



Schema di architettura con unità di canale

Le unità di canale non realizzano un semplice collegamento fisico tra le varie unità; esse sono anche dotate di capacità di elaborazione autonoma: infatti, l'unità centrale del sistema rende disponibile all'unità di canale il programma da eseguire per l'input o l'output dei dati e questo programma viene svolto in modo del tutto indipendente, anche se comunque sotto il controllo dell'unità centrale.

Le architetture dei minicalcolatori e dei PC sono in generale basate sul collegamento dei diversi componenti mediante gruppi di conduttori che si chiamano

bus: questi possono avere **parallelismo** diverso e possono essere specializzati per la trasmissione di soli dati o di soli indirizzi o di soli comandi.

Le architetture attualmente in uso sono solitamente basate sull'uso dei cosiddetti **microprocessori**, cioè di componenti uniche nelle quali sono integrate tutte le funzioni svolte dall'unità di governo e dall'unità aritmetica e logica. Ad ogni microprocessore viene associata una memoria a sola lettura (**ROM**) che contiene i microprogrammi, una memoria di lavoro (**RAM**) e uno o più dispositivi di interfaccia con le **unità periferiche**.

I microprocessori attualmente sul mercato hanno un parallelismo (o una parola) che va da 8 a 16 a 32 a 64 bit; inoltre, si ricorre sempre di più a PC che prevedono l'impiego di diversi microprocessori, ciascuno dei quali è dedicato a compiti specializzati, definiti mediante microprogrammi contenuti nelle rispettive ROM.

Autore: **Sandro Petrizzelli**

e-mail: sandry@iol.it

sito personale: <http://users.iol.it/sandry>