

Elaboratori elettronici: orientamenti e prospettive

L'evoluzione dell'elaboratore prosegue senza soste da venti anni. In questo articolo si traccia un panorama delle probabili innovazioni tecniche che caratterizzeranno gli elaboratori degli anni '70

di Franco Filippazzi

L'elaboratore elettronico è ormai un protagonista di questa nostra società tecnologica. Lo troviamo nelle industrie, negli ospedali, sulle astronavi, nella scuola. Ci sono oggi nel mondo decine di migliaia di impianti, per un valore complessivo vertiginoso. Eppure soltanto 20 anni fa esso in pratica non esisteva. Ciò che colpisce nella storia dell'elaboratore non è però soltanto la rapidità con cui si è affermato e diffuso, ma anche il ritmo incessante, per non dire frenetico, della sua evoluzione tecnica. In un arco di tempo relativamente breve si contano infatti già tre successive « generazioni » di macchine profondamente diverse tra loro, (si veda la figura a pag. 41), e ormai si è in vista di una ulteriore, quarta generazione.

Se tralasciamo la preistoria dell'elaborazione automatica delle informazioni, la prima generazione abbraccia all'incirca il periodo degli anni '50. L'elemento tecnologico che caratterizza le macchine di questo periodo è il tubo elettronico, come dispositivo di commutazione e amplificazione. Le macchine sono mastodontiche in rapporto alle prestazioni e dissipano in calore grandi quantità di energia. Esse si guastano inoltre con molta facilità. Il software (con questo termine si indica l'insieme di tutto ciò che rende più facile l'uso dell'elaboratore) di questo periodo è estremamente limitato. La programmazione è molto difficile perché si usa il linguaggio interno della macchina o linguaggi simbolici a basso livello. L'efficienza delle macchine di questa generazione è inoltre limitata dagli organi di ingresso e uscita; la scheda perforata è il principale supporto dei programmi e dei dati.

La seconda generazione nasce sul finire degli anni '50 con l'avvento dei dispositivi elettronici a stato solido. Il transistor si sostituisce al tubo elettro-

nico con enormi vantaggi. È drasticamente più piccolo, dissipa pochissima energia, permette velocità di elaborazione maggiori, ed è inoltre di gran lunga più sicuro. Si passa infatti di colpo dalla vita media di 10 000 ore del tubo a oltre un milione di ore col transistor. Ciò permette di aumentare notevolmente la complessità e la potenza delle macchine. Contemporaneamente il software riceve un grande impulso. Compiono i linguaggi simbolici ad alto livello, che permettono di sottoporre i problemi alla macchina in termini molto vicini al linguaggio umano. Alcuni di questi linguaggi, come il COBOL e il FORTRAN, riscuotono un grande successo; essi diverranno i più importanti linguaggi della generazione successiva. La dotazione di unità periferiche si arricchisce considerevolmente in questo periodo; il nastro magnetico costituisce il supporto caratteristico della seconda generazione. Il calcolatore comincia a essere meglio utilizzato anche mediante l'introduzione di una parziale simultaneità delle operazioni (calcolo interno sovrapposto a operazioni periferiche). La gestione del sistema, prima affidata completamente all'operatore, è ora in parte controllata automaticamente dal software. Infine, la maggiore complessità delle macchine spinge a studi sistematici sui metodi per l'individuazione automatica dei guasti.

La terza generazione appare nella seconda metà degli anni '60 ed è contrassegnata dalla tecnica dei dispositivi integrati. All'incirca nello stesso spazio occupato da un transistor, abbiamo ora una funzione logica completa. I vantaggi già acquisiti col transistor (velocità, dissipazione di energia, dimensioni, costo, durata) risultano ulteriormente esaltati. In questa generazione vengono inoltre introdotti nuovi concetti di operazione. Tra questi, fon-

damentale, l'elaborazione a distanza, mediante la trasmissione dei dati sulle linee telefoniche o con altri mezzi. Si afferma così l'operazione in « tempo reale », in cui l'elaboratore interagisce con l'utente, anche se lontano, all'atto stesso in cui presso quest'ultimo nasce la necessità. Non direttamente connesso con l'esigenza di cui sopra, ma mirante soprattutto a migliorare l'utilizzazione del calcolatore è il funzionamento « a partizione di tempo » (*time-sharing*), in cui più utenti, da terminali diversi, utilizzano - in pratica contemporaneamente - uno stesso elaboratore. Si generalizza inoltre in questo periodo la sovrapposizione delle operazioni (« multiprogrammazione »), introdotta nella seconda generazione solo sui calcolatori di grandi dimensioni. Il parco delle unità periferiche è ora molto attrezzato; in questo campo il disco magnetico assume una posizione fondamentale. Data la complessità dei sistemi della terza generazione, si impone una loro gestione automatica; questo compito è devoluto al software tramite i cosiddetti « sistemi operativi ». Per motivi analoghi, si mettono in atto metodologie che permettono alla macchina di diagnosticare da sé la maggior parte dei suoi possibili guasti. In questo periodo infine, viene messo l'accento sul concetto di « famiglia » o « linea » di elaboratori tra loro compatibili. Questo concetto nasce dall'esigenza di offrire all'utente la possibilità di aumentare la sua capacità di elaborazione sostituendo solo una parte del sistema già in suo possesso.

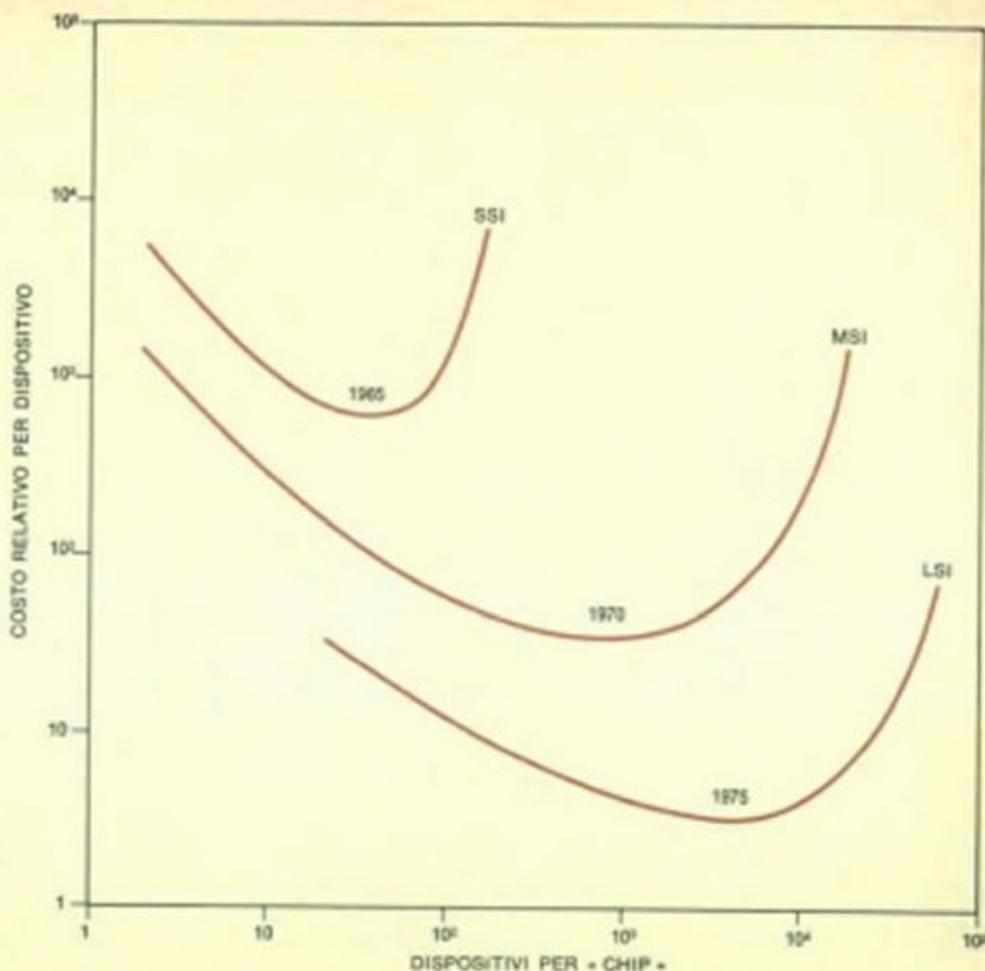
La terza generazione

La terza generazione, attualmente in corso, pur rappresentando un formidabile progresso rispetto al passato, è da considerare però ancora una tappa intermedia nell'evoluzione dell'elaborato-

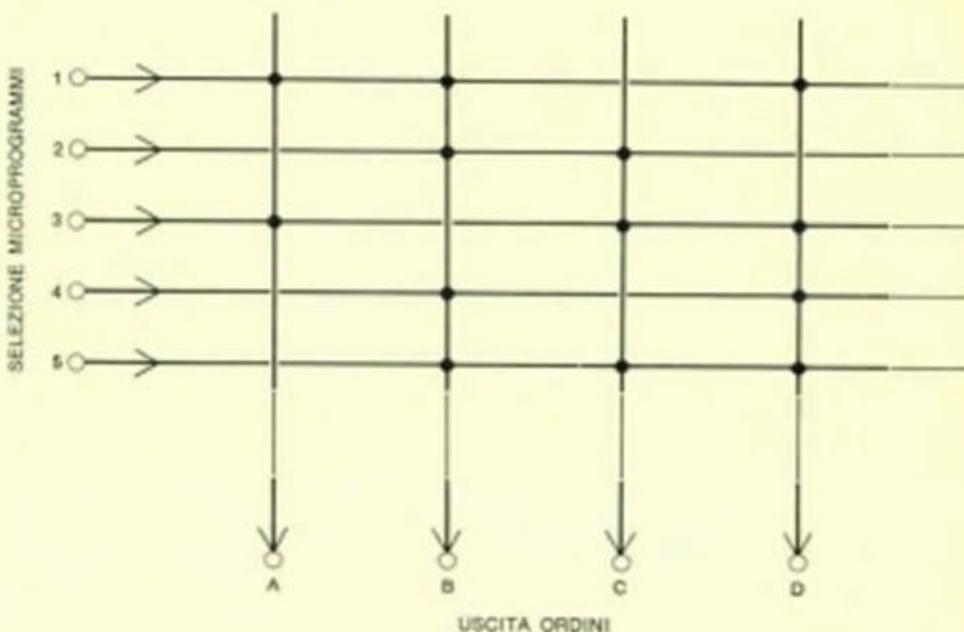
re. In essa sono infatti presenti alcune limitazioni e contraddizioni il cui superamento non può essere ottenuto mediante perfezionamenti nell'ambito delle concezioni attuali.

Dal punto di vista tecnologico, la situazione presenta una serie di squilibri e di vincoli. Uno degli squilibri (anche se non peculiare di questa generazione) è la differenza di velocità tra dispositivi logici e dispositivi di memoria. I circuiti logici attuali commutano tra i loro due stati in milionesimi di secondo, mentre le memorie a nuclei magnetici, che ancora costituiscono la maggioranza delle memorie « interne », operano in tempi nell'ordine del milionesimo di secondo. Il divario diventa enorme se si passa alle memorie « esterne », che hanno tempi operativi dell'ordine dei centesimi di secondo (dischi) o addirittura dei secondi (nastri). Anche se sono stati introdotti artifici ingegnosi per migliorare le cose, nessuno di essi risolve in modo razionale il problema, e l'architettura dei sistemi attuali rimane influenzata in modo fondamentale da questi limiti della tecnologia.

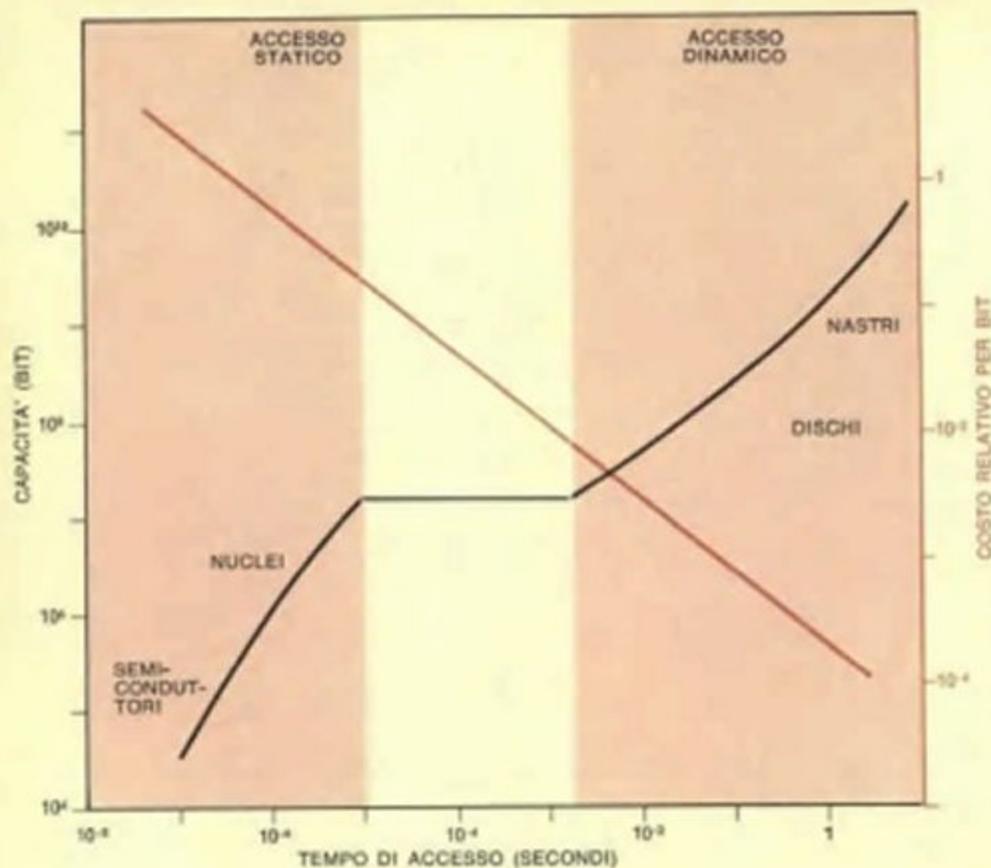
Un altro aspetto caratteristico è costituito dal crescente squilibrio tra l'*hardware* (con questo termine si indica l'insieme delle apparecchiature che costituiscono l'elaboratore) e il software, in termini di complessità, costo e tempi di sviluppo. Sotto questo aspetto, la situazione nella terza generazione si è completamente rovesciata rispetto a quella esistente nella prima generazione; mentre infatti nei primi calcolatori il peso dell'*hardware* era assolutamente preponderante, successivamente il peso del software è venuto progressivamente aumentando fino a eguagliare, se non a superare ormai, quello dell'*hardware*. Questa evoluzione è dovuta alla necessità di semplificare l'uso della macchina fornendo all'utente una serie di procedure che lo svincolino dall'obbligo di conoscere la logica interna e il modo di procedere della macchina stessa (software base); inoltre è prevalso l'uso di fornire un elaboratore con un corredo di programmi ottimizzati per la particolare classe di problemi che l'utente deve risolvere (software applicativo). Estrapolando nel futuro la tendenza in atto per quanto concerne il software, si va incontro a una situazione non accettabile per ragioni di risorse umane e tempi di sviluppo. Occorre quindi orientare la progettazione del sistema in modo da semplificare la struttura del software, delegando all'*hardware* un crescente numero di procedure di elaborazione e di funzioni algoritmiche. Dovrà cioè diminuire la di-



In un circuito integrato il costo per dispositivo dipende dal livello di complessità del circuito stesso, ossia dal numero di dispositivi realizzati in uno stesso blocchetto (*chip*) di materiale. Nei circuiti integrati usati in pratica, la complessità si aggira attorno al valore che rende minimo il costo. Tale valore si sposta col progredire della tecnologia, e quindi col passare degli anni, dallo stadio di bassa integrazione (SSI), a quello attuale di media integrazione (MSI), e più in basso con l'integrazione su larga scala (LSI).



Schema di principio di una « memoria di controllo ». Inviando un segnale elettrico in una delle linee orizzontali, esso si trasmette su tutte le linee verticali che risultano a essa connesse. Per esempio, attivando la linea 3, si ottiene la configurazione di uscita (in linguaggio binario) $a = SI, b = NO, c = SI, d = SI$. A ogni microprogramma corrisponde quindi una serie di « ordini » che controlla l'operazione dell'elaboratore.



Panorama delle memorie attuali. Esso mostra l'interdipendenza dei parametri velocità, capacità (in nero) e costo (in colore), e le zone di impiego delle diverse tecnologie.

stanza oggi esistente tra il linguaggio proprio della macchina e i linguaggi di programmazione, incorporando nel primo maggiori contenuti procedurali e complessità funzionali.

Un ulteriore aspetto limitativo è co-

stituito dalla notevole rigidità dei sistemi attuali nei confronti delle specifiche esigenze dell'utente. Ciò si traduce in definitiva in una non ottimizzazione del rapporto costo/prestazioni per l'utente stesso, che viene assogget-

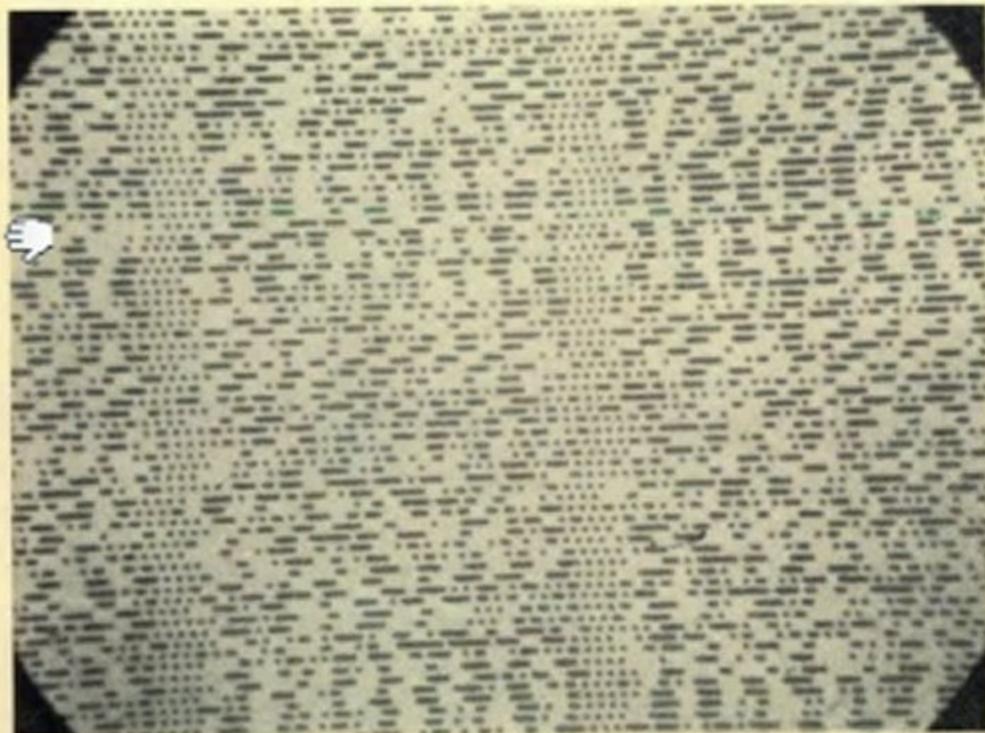
tato ad appesantimenti di hardware e di software a lui non necessari. Il time-sharing rappresenta una certa soluzione per questo tipo di problema, in quanto ripartisce lo stesso calcolatore tra più utenti, in modo da sfruttare meglio la potenzialità. Questo modo di utilizzare l'elaboratore è indubbiamente un orientamento fondamentale per gli anni a venire; esso non esclude però l'impiego di elaboratori locali, soprattutto di piccole-medie dimensioni, per i quali deve essere prevista una ottimizzazione economica migliore di quella attuale. In quest'ultimo caso occorre trovare una soluzione che consenta di «modellare» l'elaboratore sulle necessità dell'utente pur facendo uso di un hardware sostanzialmente standardizzato. In altri termini si tratta di superare una antinomia del tipo specializzazione-standardizzazione.

In generale, si può dire che lo sfruttamento della potenza di calcolo oggi esistente nel mondo lascia ancora molto a desiderare. Occorre rendere più facile e naturale l'interazione dell'uomo con la macchina. Bisogna migliorare la «disponibilità» dei sistemi, riducendo la frequenza dei guasti e i tempi di manutenzione. Sarebbe inoltre molto attraente realizzare l'associazione di calcolatori, anche molto distanti tra loro, in una specie di sistema di vasi comunicanti, in cui le risorse globali risultino a disposizione di ogni associato.

I problemi di fondo accennati non trovano soluzione pratica nell'ambito di perfezionamenti dei sistemi attuali, che stanno ormai raggiungendo i loro limiti; si richiede invece l'introduzione di concetti nuovi, della portata di quelli che in passato hanno individuato l'avvento di una nuova generazione di elaboratori. Quali sono gli elementi innovativi o comunque caratterizzanti che troveremo nei sistemi degli anni '70? Per quanto ci è consentito di prevedere oggi, si possono individuare alcuni punti fondamentali che toccano la progettazione dell'elaboratore nei suoi vari aspetti: dalla introduzione delle tecniche di integrazione su larga scala, all'impiego sistematico del *firmware*, dalla comparsa di memorie di massa di nuova concezione, alla realizzazione di reti di calcolatori, dalla autodiagnostica e autoriparazione dei guasti, alla introduzione di nuovi mezzi per il «colloquio» uomo-macchina.

L'integrazione su larga scala

L'integrazione su larga scala (LSI = *Large Scale Integration*) è sostanzialmente l'evoluzione e il perfeziona-



Microfotografia di una porzione di memoria ottica che mostra i dati registrati in codice binario mediante evaporazione selettiva del substrato ottenuta con un fascio laser.

mento delle tecniche di integrazione circuitale introdotte con la terza generazione di elaboratori. Dalle poche decine di dispositivi per blocchetto di silicio (*chip*) di qualche anno fa, si è oggi arrivati a diverse centinaia di dispositivi. La LSI ci prospetta ora complessità che sono dieci, cento o più volte maggiori di quelle attuali.

Volendo fare una previsione sulla complessità dei circuiti integrati che verranno usati nei prossimi anni, dobbiamo tener conto della correlazione esistente tra complessità, ossia numero di dispositivi per *chip*, e costo unitario (si veda la figura in alto a pagina 39). Per ogni epoca esiste un valore ottimale di complessità; sotto questo valore il costo aumenta perché non si sfruttano le capacità della tecnologia; sopra, perché aumentano gli scarti di produzione. La complessità dei circuiti effettivamente usati non è quella massima che la tecnologia dell'epoca sarebbe in grado di realizzare, ma quella che ottimizza il costo. (Questo criterio può non essere valido per applicazioni speciali, come quelle spaziali, dove i requisiti funzionali precedono, di regola, le considerazioni di costo).

Quali vantaggi comporterà la LSI per gli elaboratori? Oltre alla evidente riduzione di ingombro - ma in buona parte proprio a causa di questo - la LSI significa anche maggiore velocità e minore dissipazione di energia. Si dovrebbe inoltre ottenere un decremento del tasso di guasti, a causa della drastica riduzione del numero di contatti e di saldature, che sono notoriamente tra le principali cause di malfunzionamento delle apparecchiature elettroniche. Infine la LSI significa importanti vantaggi economici. Ciò è dovuto sostanzialmente al fatto che la fabbricazione dei circuiti integrati avviene mediante processi automatici che permettono di realizzare contemporaneamente un numero elevatissimo di dispositivi (*batch process*).

La domanda che ci poniamo ora è: in quale modo specifico potranno essere sfruttati negli elaboratori i vantaggi funzionali ed economici offerti dalla LSI? O, in termini più precisi: costituisce la LSI una «chiave tecnologica» per realizzare delle soluzioni altrimenti non possibili in pratica.

Un primo tema che si può intravedere sotto questo profilo riguarda il superamento della antinomia specializzazione-standardizzazione cui abbiamo già accennato, allo scopo di ottimizzare il rapporto costo/prestazioni degli elaboratori. Una soluzione del problema consiste nel progettare l'elabora-

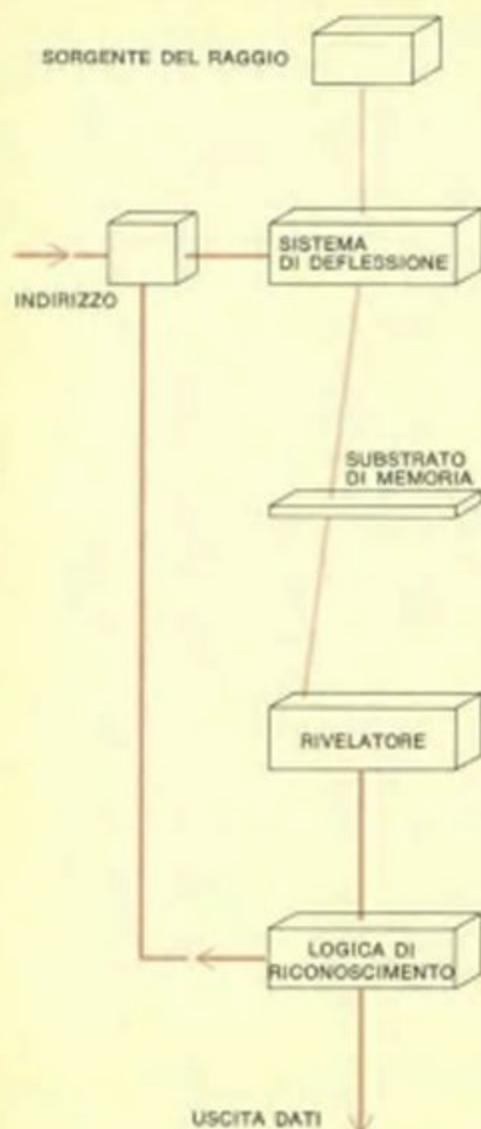
HARDWARE						
GENERAZIONE	ELEMENTO CARATTERISTICO	PARAMETRI TIPICI RELATIVI				
		VELOCITÀ	DIMENSIONI	DISSIPAZIONE	TASSO DI GUASTI	COSTO
PRIMA	VALVOLA	1	1	1	1	1
SECONDA	TRANSISTORE	10	1/100	1/1000	1/100	1/10
TERZA	CIRCUITO INTEGRATO	100	1/1000	1/1000	1/1000	1/100

In queste tabelle sono raccolte le principali caratteristiche che hanno distinto e caratterizzato le tre «generazioni» nelle quali si usano distinguere i calcolatori elettronici finora esistenti. Dalle tabelle risulta evidente come la distinzione in «generazioni» sia piuttosto approssimativa e costituisca soltanto un utile sistema di riferimento. (I parametri tipici relativi danno gli ordini di grandezza riferiti a una funzione logica elementare).

SOFTWARE E SISTEMISTICA						
GENERAZIONE	LINGUAGGI DI PROGRAMMAZIONE	MODO OPERATIVO		SUPPORTO CARATTERISTICO	GESTIONE DEL SISTEMA	DIAGNOSTICA
		LOCALE	SERIALE (uniprogrammazione)			
PRIMA	LINGUAGGIO MACCHINA O SIMILE	LOCALE	SERIALE (uniprogrammazione)	SCHEDA PERFORATA	OPERATORE	MANUALE
SECONDA	COMPARSA DEI LINGUAGGI AD ALTO LIVELLO	LOCALE	PARZIALMENTE SIMULTANEO	NASTRO MAGNETICO	OPERATORE + SOFTWARE	AIUTATA DALLA MACCHINA
TERZA	LINGUAGGI AD ALTO LIVELLO LINGUAGGI CONVERSAZIONALI	LOCALE E A DISTANZA REAL TIME TIME SHARING	SIMULTANEO (multiprogrammazione)	DISCO MAGNETICO	SOFTWARE (sistemi operativi)	SOSTANZIALMENTE AUTOMATICA

tore in modo che le parti che lo compongono siano nella grande maggioranza indipendenti dall'impiego specifico della macchina e che questa si possa predisporre per un certo tipo di applicazioni mediante l'inserimento di un blocco particolare, di dimensioni e costo molto ridotti rispetto al resto del sistema. Questo blocco contiene, in sostanza, la logica di controllo dell'intera macchina sotto forma di «microprogrammi», ognuno dei quali corrisponde a una operazione elementare della macchina stessa. Una qualsiasi elaborazione è realizzata mediante una opportuna sequenza di microprogrammi. Il calcolatore viene specializzato per i vari ruoli per cui è previsto, attraverso il cambiamento del suo patrimonio di microprogrammi, cioè del suo cosiddetto firmware.

In effetti, il principio della microprogrammazione non è affatto nuovo, ma è stato finora sfruttato solo parzialmente soprattutto perché non era disponibile una adeguata soluzione hardware per realizzarlo. La LSI offre ora la possibilità di una soluzione funzionale ed economica del problema. I microprogrammi vengono realizzati mediante una struttura semplice e regolare come il reticolo o matrice rappresentato nell'illustrazione in basso a pagina 39. In questo schema, le informazioni sono rappresentate mediante connessioni selettive ai nodi del reticolo. Questo organo si può considerare una forma particolare di memoria («memoria di controllo»), in cui le informazioni immagazzinate sono fisse. Essa viene perciò chiamata anche «memoria a sola lettura» (*Read-Only Store*, nella termi-



Schema di principio del funzionamento di una memoria a raggio energetico. Il raggio generato da un laser (raggio ottico) o da un cannone elettronico (raggio di elettroni) viene deviato sul punto desiderato del substrato di memoria tramite un sistema di deflessione al quale viene presentato l'indirizzo del punto stesso. Il substrato di memoria si può considerare un insieme di celle di memoria in ognuna delle quali — una cella alla volta — il raggio registra o legge un'informazione elementare (bit). In generale, sul substrato, oltre ai dati veri e propri, sono registrati anche gli indirizzi di blocchi di bit e altre informazioni atte a dirigere il raggio nei punti desiderati. Mediante queste informazioni, opportunamente discriminate, si può retroagire sul sistema di deflessione per correggere la direzione del raggio. Le densità di informazione raggiungibili con memorie di questo tipo sono estremamente elevate. Su un millimetro quadrato del substrato si possono infatti immagazzinare fino a diecimila bit usando un raggio ottico e a oltre un milione di bit usando invece un raggio elettronico.

nologia anglosassone). In un altro tipo di realizzazione, simile alla precedente ma più flessibile, le informazioni possono essere modificate, seppure con velocità assai inferiore a quella con cui

possono essere lette (*Read-Mostly Store*). Per questo tipo di memoria di controllo si prospettano diverse soluzioni tecnologiche, tutte ancora in fase di sviluppo. E infine da osservare che la funzione in oggetto può essere realizzata mediante una memoria vera e propria, cioè modificabile senza restrizioni. Ragioni funzionali (velocità, sicurezza) e di costo, fanno preferire le soluzioni precedenti.

Il firmware, inteso come impiego sistematico e generalizzato della microprogrammazione, costituirà certamente uno degli aspetti fondamentali e caratteristici della prossima generazione di elaboratori. Esso si porrà come intermediario tra un hardware sostanzialmente standardizzato e un software sempre più rivolto alle applicazioni specifiche. Il firmware avrà perciò una influenza su entrambe queste componenti dell'elaborazione dei dati, e potrà contribuire a modificare lo scorporo in atto tra hardware e software cui abbiamo in precedenza accennato, assumendo parte delle funzioni attualmente svolte dal software. Infatti, tramite il firmware, uno stesso modello di hardware può apparire al programmatore dotato di un qualsiasi linguaggio macchina. Al limite, in teoria, si potrebbe fare un firmware così potente da far diventare linguaggio di macchina gli stessi linguaggi simbolici ad alto livello. Analogamente, molte funzioni svolte dai « sistemi operativi », oggi parte tipica del software di base, verranno probabilmente realizzate in firmware. Il firmware costituirà anche lo strumento appropriato per ottenere la compatibilità tra macchine diverse; esso permette infatti all'utente che passa da un modello a un altro, di « simulare » sul nuovo elaboratore quello precedente, in modo da poter continuare ad usare i programmi già impiegati. In definitiva, si può prevedere che il firmware prenderà in carico molte delle funzioni attuali del software di base, modificando una tendenza non ottimale oggi in atto, e conferendo flessibilità nuove agli elaboratori.

Le memorie

Una funzione fondamentale in qualsiasi processo di elaborazione delle informazioni è notoriamente quella di memoria. La memoria ideale è poco costosa, di piccole dimensioni, di grande capacità; consuma poca energia e opera alla stessa velocità della logica del calcolatore. Oggi non esiste nessuna tecnologia che realizzi una memoria così fatta. Esistono invece diverse tecnologie, ognuna delle quali ottimizza

solo una parte dei requisiti della memoria ideale. I progettisti di elaboratori hanno pertanto suddiviso la memoria in una serie di memorie, ognuna delle quali assolve a compiti specializzati. In ogni sistema di elaborazione si ritrova quindi una « gerarchia » di memorie, corrispondenti a fasce diverse di prestazioni e di costo. In generale, quanto più una memoria è veloce, tanto più è costosa e, soprattutto per quest'ultimo motivo, tanto più piccola è la sua capacità, ossia i bit che in essa vengono immagazzinati (si veda la figura in alto a pag. 40). Nella gerarchia delle memorie intervengono anche ulteriori criteri di classificazione oltre a velocità, capacità e costo. Un criterio si riferisce al tipo di accesso, cioè al modo con cui si arriva alla cella di memoria desiderata; le memorie sono pertanto suddivise in statiche e dinamiche, a seconda che l'accesso avvenga con mezzi puramente elettronici oppure richieda movimenti meccanici di parti della memoria. Nelle memorie statiche l'accesso è normalmente del tipo detto « a caso », cioè si può saltare da una cella di memoria a un'altra qualsiasi senza passare attraverso celle intermedie; il contrario avviene per le memorie del secondo tipo, che presentano un accesso « seriale ». Le memorie statiche sono caratterizzate da velocità molto maggiori di quelle dinamiche, e sono pertanto usate come memorie principali, che operano col ritmo proprio della logica della macchina. Le memorie dinamiche consentono invece di immagazzinare economicamente grandi quantità di informazioni (memorie di massa), che devono però essere scambiate con la memoria principale per essere elaborate.

Quali sono le prospettive per le memorie principali? Negli ultimi 15 anni i nuclei magnetici hanno dominato questo campo delle memorie. L'anellino di ferrite — un pioniere tutt'ora sulla breccia — ha indubbiamente aperto una nuova era nella storia dell'elaboratore elettronico, dando la possibilità di realizzare memorie veloci, sicure, poco costose. Oggi la supremazia degli elementi magnetici viene posta in discussione dall'avvento della LSI. Sono le stesse motivazioni generali che abbiamo visto in precedenza (miniaturizzazione, velocità, fabbricazione con processi automatici, ecc.), che fanno della LSI un candidato anche per le applicazioni di memoria. Già oggi le memorie a semiconduttore costituiscono l'unica soluzione quando si richiedano velocità elevatissime, non ottenibili coi nuclei. Ma le prospettive delle memorie a semiconduttore sono molto più am-

biziose: esse si presentano come vantaggiosa alternativa ai nuclei magnetici su tutto l'arco delle memorie principali. La competizione tra le due tecnologie si è aperta, in pratica, da poco tempo. Il nucleo non è affatto finito; però ci sono pochi dubbi che saranno le memorie LSI ad avere il sopravvento nei prossimi anni. Infine, citiamo anche il fatto che la LSI possa costituire una soluzione conveniente per la zona scoperta del grafico di pagina 40, mediante dispositivi seriali statici (« registri a scorrimento »).

Prima di lasciare l'argomento, c'è una considerazione di carattere generale che vale la pena di mettere in rilievo, e cioè che con la LSI si realizzerà una unificazione tecnologica fino a oggi impossibile. Al posto delle varie tecnologie che ora occorre impiegare a seconda che si tratti di logica o di memoria di controllo o di memoria di lavoro, ci sarà una sola tecnologia. Non è questa una considerazione di valore accademico in quanto ogni « interfaccia » tra tecnologie diverse significa, in generale, degradazione di prestazioni e aumento di costo.

Passando alle memorie di massa, le soluzioni attuali presentano delle limitazioni che sono in sostanza legate alla modalità con cui, in esse, si accede alle informazioni. L'operazione implica infatti lo spostamento meccanico del substrato di memoria (disco magnetico, nastro magnetico), in modo da portare la parte di esso che interessa in corri-

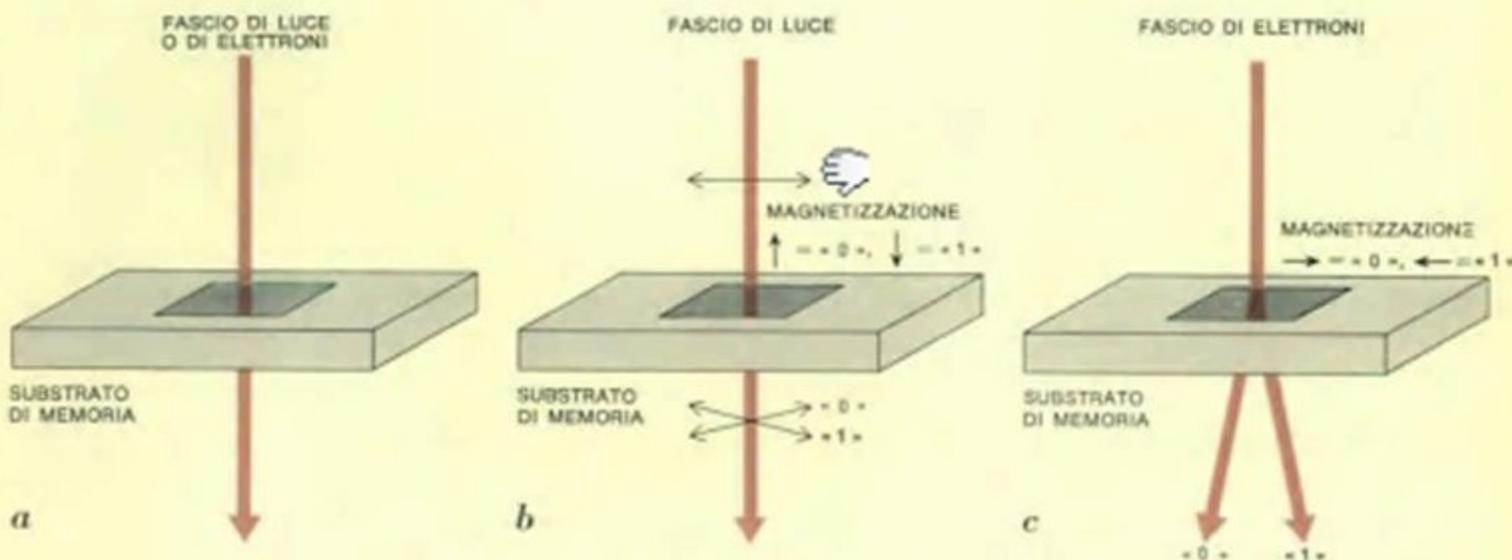
spondenza delle stazioni (testine) di lettura-scrittura. L'accesso meccanico comporta non solo fondamentali limitazioni nella velocità della memoria, ma incide anche negativamente sulla « fidatezza » del funzionamento.

Una soluzione a questo tipo di problemi è rappresentata da un sistema di memoria in cui l'accesso è realizzato mediante un sottile pennello ottico o elettronico, che viene deflesso sul substrato di memoria. In linea generale, il substrato di memoria contiene, oltre ai dati, anche informazioni atte a indirizzare correttamente il raggio. Il sistema è così in grado di conoscere la posizione raggiunta dal pennello di radiazione e può eventualmente, attraverso una rete di retroazione, correggerne la deflessione, portandolo a incidere nel punto richiesto del substrato. L'uso di un pennello di radiazione come mezzo di accesso, oltre che a conseguire velocità di operazione del tutto irrealizzabili con sistemi meccanici, permette anche di ottenere densità di informazioni sul substrato di memoria di gran lunga superiori a quelle ottenibili con i sistemi di registrazione convenzionali. Vediamo ora più da vicino su quali principi si basi il funzionamento di queste memorie. Va premesso che esistono due classi di soluzioni: l'una consente di modificare il substrato per scrivere nuove informazioni, l'altra invece è del tipo a sola lettura.

Nelle memorie del primo tipo, cioè *modificabili*, il substrato delle informa-

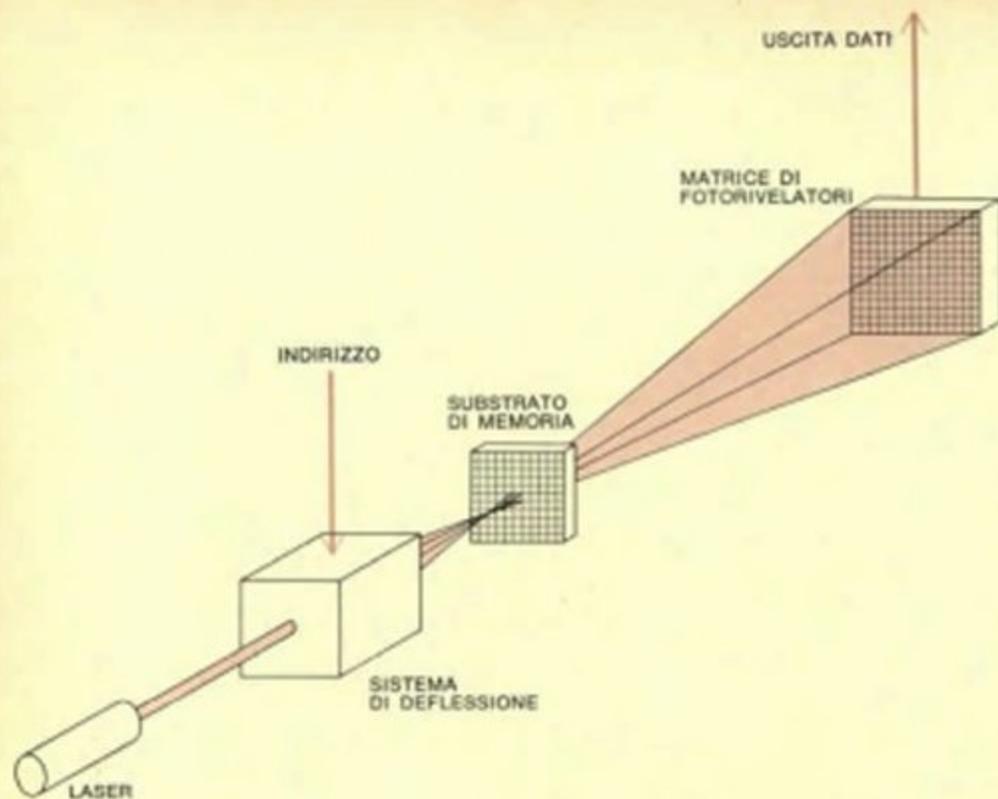
zioni è costituito da una pellicola sottile di opportuno materiale magnetico. La scrittura è basata su effetti termomagnetici, in particolare sulla variazione del campo coercitivo del materiale in funzione della temperatura (si veda la figura qui sotto). In sostanza, il metodo consente di cambiare selettivamente la magnetizzazione in un punto qualsiasi del substrato. Usando i due stati di saturazione magnetica del materiale, si può pertanto scrivere in ogni areola, una informazione binaria. La lettura è basata su effetti diversi a seconda che il pennello è di natura ottica (laser) o elettronica. Nel primo caso si sfruttano effetti magnetooptici; nel secondo caso si sfrutta la forza esercitata dal campo magnetico del substrato sugli elettroni incidenti. Ricerche su questi tipi di memoria sono in corso anche in Italia, presso il Laboratorio di ricerche della Honeywell Information Systems Italia di Pregnana Milanese.

Esistono applicazioni in cui si richiede che la memoria conservi una grande quantità di informazioni (superiore ai 10^{12} bit) senza che occorra modificarle, ma solo aggiungerne di nuove. A queste memorie è stato attribuito il nome di *memorie archivio*. Anche queste memorie possono essere convenientemente realizzate mediante sistemi ad accesso ottico o elettronico. In quelle di quest'ultimo tipo, il substrato di memoria è costituito, in una soluzione, da un sottile strato metalli-

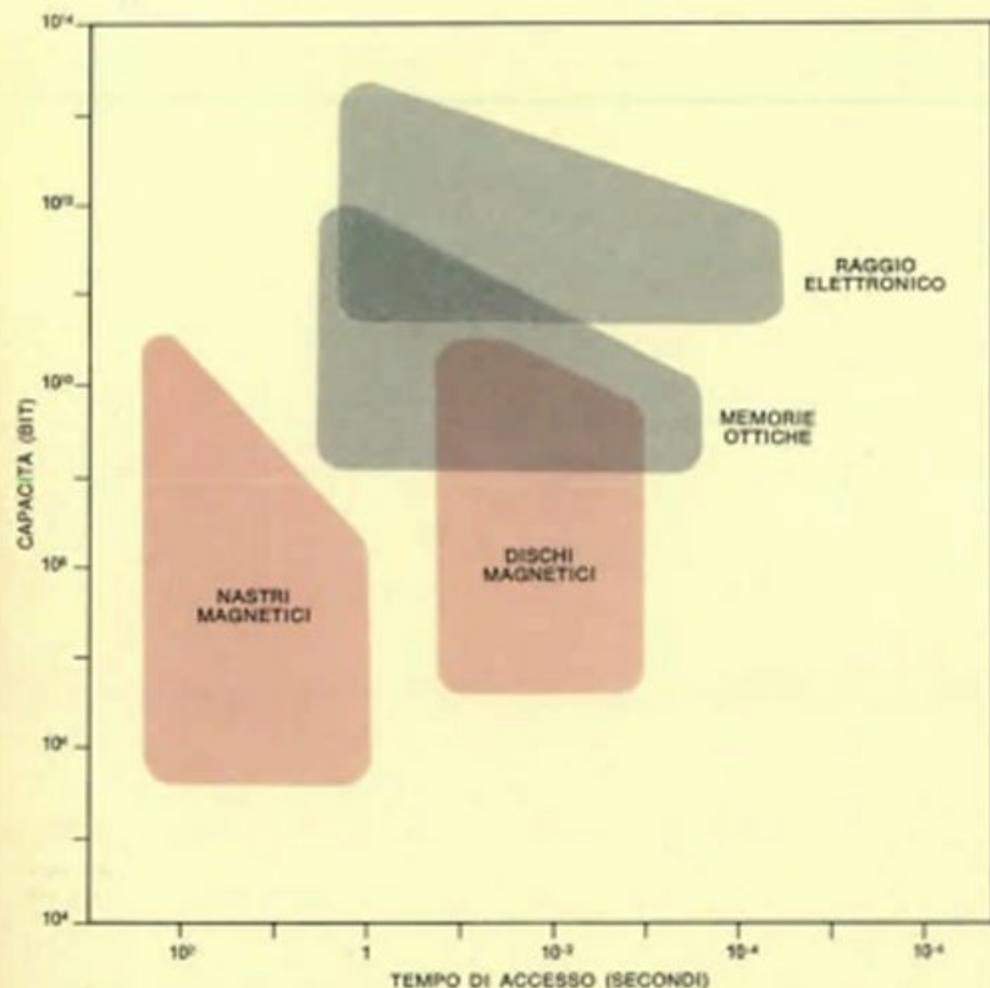


Principi di operazione delle memorie elettroottiche modificabili. Nel principio di scrittura termomagnetica (a), il fascio (ottico o elettronico) riscalda una piccola area del substrato attorno al punto di incidenza. In questa area il materiale assume un campo coercitivo inferiore a quello circostante. Applicando a tutto il substrato una eccitazione magnetica di valore intermedio, si può orientare nella direzione del campo esterno il materiale della zona riscaldata senza alterare lo stato del materiale circostante. Per i principi di lettura si hanno due casi diversi. Se

il raggio è di tipo ottico (b), si può conoscere lo stato del materiale dalla rotazione del piano di polarizzazione del raggio trasmesso (rotazione che è dell'ordine di qualche grado). Se il raggio è elettronico (c), l'informazione si ricava dalla deviazione che esso subisce attraversando il materiale (deviazione che è dell'ordine di qualche centesimo di grado). Sia con il raggio ottico che con quello elettronico la lettura può essere eseguita anche mediante i raggi riflessi dal materiale anziché mediante i raggi trasmessi, come nel caso illustrato in questa figura.



Schema di principio di memoria olografica. Sul substrato di memoria (lastra fotografica, se la memoria è a « sola lettura »), è registrata una serie di ologrammi ciascuno dei quali corrisponde a una « pagina » di bit (N ologrammi di n bit ciascuno). Il fascio ottico è diretto sull'ologramma selezionato, che viene ricostruito nel piano del rivelatore che è costituito da una matrice di elementi, tanti quanti sono i bit di una pagina. Indicativamente, un ologramma occupa 1 millimetro quadrato e contiene diecimila bit.



Previsione sulle memorie di massa nella seconda metà degli anni '70. Le prestazioni potenziali dei nuovi tipi di memorie sono qui a confronto con i sistemi convenzionali.

co. Una informazione binaria viene rappresentata mediante la presenza o l'assenza di un piccolissimo foro (circa 0,1 micron di diametro) praticato nello strato suddetto. La scrittura, cioè la realizzazione dei microfori, è fatta mediante il fascio elettronico della memoria, che fa evaporare in modo selettivo lo strato metallico. Lo stesso fascio elettronico (con intensità ridotta) è usato per l'operazione di lettura: esso passa attraverso il substrato e raggiunge il rivelatore solo dove trova un microforo. Nelle memorie archivio di tipo ottico, il substrato di memoria è invece costituito da una lastra fotografica. Le informazioni binarie vengono rappresentate sulla lastra mediante areole trasparenti od opache alla luce, ogni areola corrispondendo ad un bit.

Una classe interessante di memorie ad accesso ottico è quella delle memorie olografiche (si veda la figura in alto a sinistra). In questo caso, il substrato di memoria non immagazzina configurazioni di bit, ma il loro ologramma. Più precisamente, sulla lastra sono registrati una serie di ologrammi, ognuno dei quali corrisponde non a un bit, ma a un insieme, o « pagina », di bit. Indirizzando il raggio laser sull'ologramma prescelto, questo viene ricostruito in corrispondenza del rivelatore. Quest'ultimo è costituito da un mosaico di elementi fotoelettrici, che converte le informazioni ottiche in segnali elettrici utilizzabili dall'elaboratore. Le memorie olografiche presentano diversi pregi. Anzitutto l'indirizzamento alla « pagina » consente una drastica riduzione della complessità di deflessione del raggio laser (punto debole delle memorie ottiche) in confronto al caso dell'indirizzamento al bit. La selezione di una intera pagina per volta, ossia di un largo numero di bit contemporaneamente, consente inoltre la possibilità di elevate velocità di lettura. Infine, per le caratteristiche di ridondanza proprie della olografia, diminuiscono le criticità di posizionamento del fascio ottico, e vengono neutralizzati in larghissima misura gli effetti della polvere e dei graffi sulla pellicola di memoria. Esistono però anche aspetti meno positivi e limitazioni, in particolare per quanto riguarda la scrittura e la modificabilità delle informazioni.

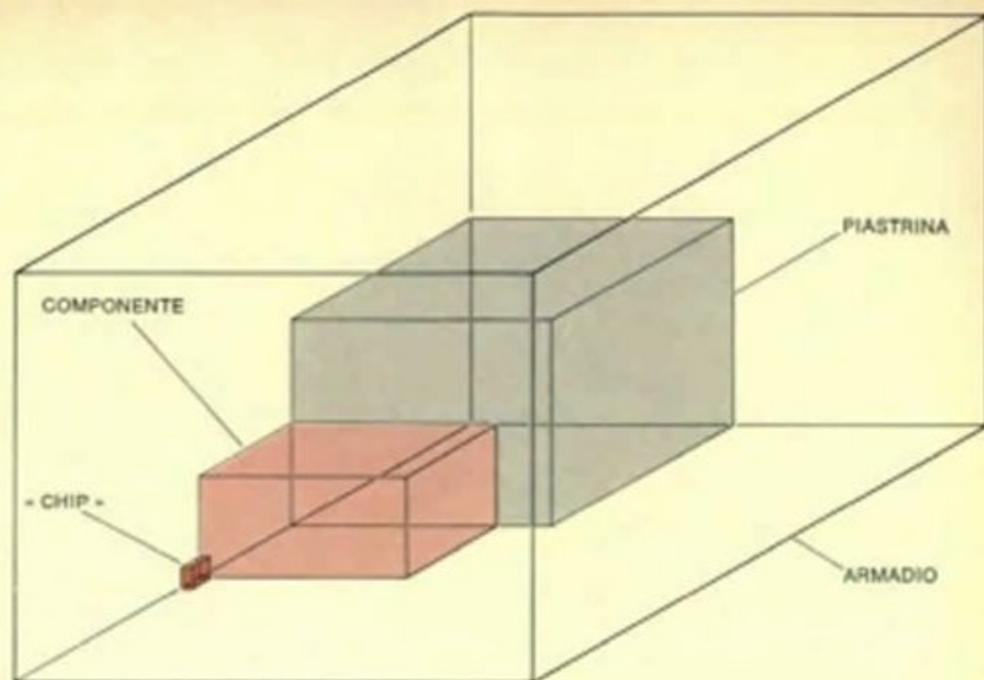
È molto probabile che durante gli anni '70 qualcuna delle nuove memorie illustrate in questo paragrafo esca dai laboratori di ricerca e venga effettivamente utilizzata. Se si confrontano i due approcci - quello ottico e quello elettronico - si trovano pro e contro per ciascuno di essi. Le memorie a fascio elettronico richiedono che il sub-

strato operi nel vuoto anziché nell'aria come quelle ottiche; però la deflessione è molto più facile e la densità delle informazioni assai maggiore (superiore a 10^4 bit/cm² nelle elettroniche, contro circa 10^6 bit/cm² in quelle ottiche). I due tipi di memoria non vanno però visti in alternativa, essendo probabilmente diversi i loro campi di impiego. Dal punto di vista applicativo, la disponibilità di queste nuove memorie di massa avrà un profondo effetto specie sui grandi sistemi in tempo reale e *time-sharing* e potrà condurre a concezioni nuove nell'architettura degli elaboratori.

Il problema delle interconnessioni

Quando si parla di tecnologia degli elaboratori, un grande tema d'obbligo è la miniaturizzazione dei dispositivi logici e di memoria. Molto meno si parla invece dei modi per collegare tutti questi dispositivi tra di loro, in modo da farne un insieme funzionante. È questo un argomento meno affascinante per il profano, ma non per questo meno importante degli altri. Per rendercene conto si osservi l'illustrazione qui a destra. In un apparato elettronico si trova normalmente una « gerarchia strutturale », che trae origine da necessità pratiche di diversa natura: facilità di montaggio, di collaudo, di trasporto, di manutenzione e altre ancora. Comunque sia, dal blocchetto di silicio che costituisce la parte attiva della macchina, all'apparato finito, esiste una progressiva degradazione della « efficienza volumetrica »: il volume della macchina è infatti costituito per oltre il 99,99 % da interconnessioni! La cosa si riflette, se pure in una misura molto meno drammatica, anche sull'aspetto economico: le interconnessioni costituiscono infatti una parte assai rilevante del costo della macchina. Questa situazione, pur se giustificata da validi motivi pratici, non è certo entusiasmante. Gli sforzi per la miniaturizzazione a livello *chip*, i vantaggi della LSI, vengono in buona parte vanificati dalla gerarchia delle interconnessioni.

D'altronde è noto che miniaturizzare corrisponde a precise esigenze. Questo è vero non solo nelle applicazioni spaziali, dove ogni grammo di « carico pagante » ha un prezzo altissimo, ma anche nelle applicazioni meno raffinate, come i calcolatori di uso normale. L'efficienza, la velocità e la sicurezza di una macchina vanno infatti in direzione opposta alle sue dimensioni. I segnali, anche nei calcolatori, non possono viaggiare a velocità maggiore di quella della luce: sfortunatamente, an-



In questo disegno si è tentato di dare un'idea della « gerarchia strutturale » che, pur dettata da valide ragioni di ordine pratico, comporta tuttavia una enorme « degradazione volumetrica ». Il rapporto tra volume e volume attivo aumenta 500 volte nel passaggio dal « chip » LSI (rappresentato dal parallelepipedo in colore pieno) al componente (in colore tenue). Un altro aumento di 5 volte si ha passando da questo alla piastra (in grigio). Un ulteriore aumento di dieci volte si ha passando da questa all'armadio (parallelepipedo grande). Si ha quindi mediamente una degradazione totale di ben 25 000 volte; in altri termini, i « chip » di semiconduttore — la parte attiva dell'insieme — costituiscono solo una minima parte del volume di una apparecchiatura elettronica. Più del 99,99 per cento di quest'ultima sono strutture di interconnessione.

zi, sono parecchio più lenti; la dissipazione di energia e il numero di saldature sono affetti negativamente dalle dimensioni, e così via per altri aspetti. Un tema di fondo della tecnologia è quindi di cercare delle soluzioni per migliorare lo stato di cose descritto.

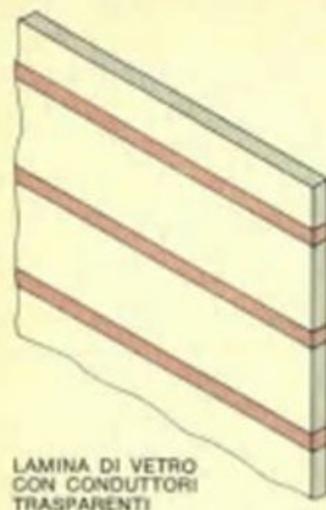
Una prima soluzione è di realizzare il maggior numero possibile di interconnessioni a livello *chip*, attuando cioè una « super LSI ». Rientra in questo quadro l'idea del calcolatore sul wafer di silicio. Anche se l'integrazione sempre più spinta è una direzione fondamentale della tecnologia, essa presenta comunque i suoi limiti, che risiedono soprattutto nella resa di fabbricazione, e quindi nel costo. È da segnalare a questo proposito che una alternativa per aumentare la complessità dei circuiti integrati è di accettare *chip* in cui una parte dei dispositivi siano difettosi. Questo metodo consiste nel predisporre nel *chip* un numero di blocchi funzionali maggiore di quello necessario per realizzare le funzioni volute (ridondanza), e nell'ottenere il circuito desiderato collegando tra loro soltanto blocchi che risultino buoni. In questo caso le interconnessioni sul silicio sono determinate sulla base di un collaudo intermedio dei dispositivi, sono cioè fatte a posteriori. Perché questo approccio abbia senso pratico, occorre che il

collaudo dei dispositivi, la scelta delle interconnessioni e la loro realizzazione, vengano fatte in modo automatico, tramite calcolatori e strumentazioni particolari. Attualmente questa soluzione è in fase di laboratorio e la sua validità economica deve ancora essere dimostrata.

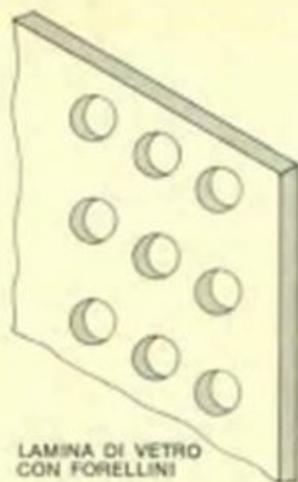
Una seconda soluzione del problema strutturale consiste nell'inserire più *chip* in uno stesso contenitore, realizzando all'interno di questo le relative interconnessioni. Si realizzano in questo modo dei macrocomponenti, che vengono poi montati su piastre convenzionali. Un macrocomponente è costituito in sostanza da un substrato ceramico, sul quale vengono realizzate le interconnessioni tra i *chip* usando tecniche che consentono una elevata miniaturizzazione delle interconnessioni medesime.

Una terza soluzione del problema è di realizzare le interconnessioni anziché mediante fili o percorsi conduttori, direttamente tramite pennelli di energia. Abbiamo visto un esempio di questo modo di fare le interconnessioni parlando delle nuove memorie di massa. È questa una tecnica efficace in applicazioni del tipo citato, ma che difficilmente potrà essere generalizzata.

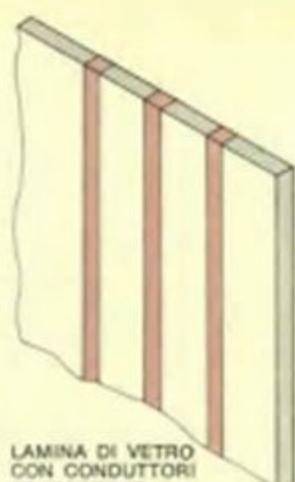
Per chiudere il tema delle interconnessioni, vale la pena di accennare a una soluzione affascinante, anche se av-



LAMINA DI VETRO
CON CONDUTTORI
TRASARENTI



LAMINA DI VETRO
CON FORELLINI



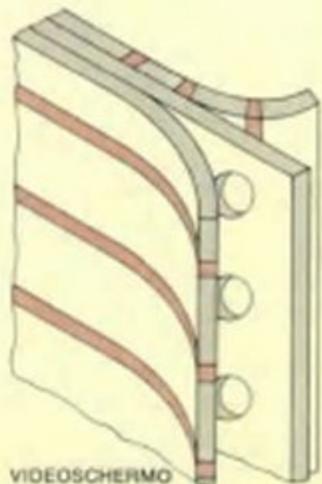
LAMINA DI VETRO
CON CONDUTTORI
TRASARENTI

Il videoschermo « a plasma » consiste sostanzialmente di tre lastre di vetro che vengono unite a sandwich. La lastra di mezzo è forellata, mentre su quelle laterali sono depositati due gruppi di conduttori trasparenti (in colore). Le cavità che si realizzano allorché le tre piastre vengono unite sono riempite di un gas facilmente ionizzabile, per esempio neo. Ogni cavità si trova all'incrocio di due conduttori ortogonali (che rappresentano i suoi elettrodi) e costituisce una minuscola lampadina a bagliore. Essa può essere accesa o spenta mediante impulsi elettrici applicati ai conduttori che la individuano. Questo schermo è sostanzialmente piatto, possiede intrinsecamente una capacità di memoria e consente un'elevata risoluzione.

veniristica. L'idea è nata nel corso di ricerche per realizzare un sistema elettronico che funzioni in modo simile alla materia grigia del cervello. Il problema di costruire un sistema simile al cervello umano non si esaurisce nel fabbricare i 10 miliardi di neuroni che lo costituiscono, ma richiede anche di risolvere il problema del « piano » delle loro interconnessioni. Ora l'idea consiste nel costruire una massa di neuristori, ossia di neuroni artificiali, riempiendo gli interstizi tra essi con un opportuno materiale. Quando una informazione (sotto forma di eccitazioni elettriche) viene immessa nel sistema, si stabiliscono nel suo interno delle vie preferenziali tra i neuristori. Ogni volta che una data via è attivata, le particelle di materiale interstiziale lungo il cammino vengono sempre più orientate, così che la corrente incontra una resistenza sempre minore. In questo modo è possibile « insegnare » alla massa di neuristori a scegliere certe vie piuttosto che altre, realizzando alla fine una rete di interconnessioni stabili. Il sistema « imparerebbe » cioè a « connettere », in modo analogo a quanto si ipotizza avvenga nel cervello umano.

Altre prospettive

Un aspetto caratteristico della elaborazione dei dati negli anni '70 sarà l'espansione dell'elaborazione a distanza. Questa tecnica è già oggi affermata con i sistemi in « tempo reale » e il *time-sharing*, in cui interviene la connessione elaboratore-terminale. Il passo successivo, che verrà attuato nei prossimi anni, sarà la realizzazione della connessione elaboratore-elaboratore. Si attuerà cioè il collegamento diretto tra elaboratori, anche di dimensioni equivalenti, in modo da formare delle reti di elaboratori. Lo scopo è di mettere a disposizione di ciascun utente della rete i programmi, gli archivi e le capacità di calcolo localizzate presso ciascun nodo della rete stessa. In una concezione di questo tipo si verranno a creare « programmotecche » e « banche di dati » specializzate presso ciascun nodo, che faranno tutte parte delle risorse comuni, eliminando gli attuali sprechi e duplicazioni di sforzi. Si pensi, per esempio, alle possibilità offerte dal collegamento tra i centri di calcolo delle università europee, secondo una proposta avanzata dal prof. Dadda del



VIDEOSCHERMO

Politecnico di Milano, rappresentante italiano presso la Commissione per l'Informatica della CEE. Le reti di calcolatori si prospettano come uno dei più interessanti concetti per il futuro, con campi di applicazione che vanno dalla ricerca scientifica alle banche, dal fisco alla salute pubblica.

Un punto chiave della elaborazione a distanza è costituito dalla trasmissione dei dati. Le tecniche relative avranno ampio sviluppo nei prossimi anni con l'impiego di tutti i mezzi di comunicazione disponibili. Già oggi si usano i satelliti per la connessione tra terminali posti in Italia e calcolatori installati oltre l'Atlantico; un ruolo fondamentale avranno anche le comunicazioni su filo.

Un altro punto focale degli sviluppi dei prossimi anni sarà quello della interazione tra l'uomo e la macchina. Affinché l'elaboratore veramente « arrivi su ogni scrivania » dovranno essere messe a punto soluzioni molto efficienti nell'area dei terminali, sia sotto l'aspetto hardware che quello software. Per il software si assisterà alla creazione di « linguaggi colloquiali » particolarmente efficaci e facili da usare. Per quanto riguarda l'hardware si può prevedere tutta una serie di sviluppi per aumentare la flessibilità dei terminali e diminuirne il costo.

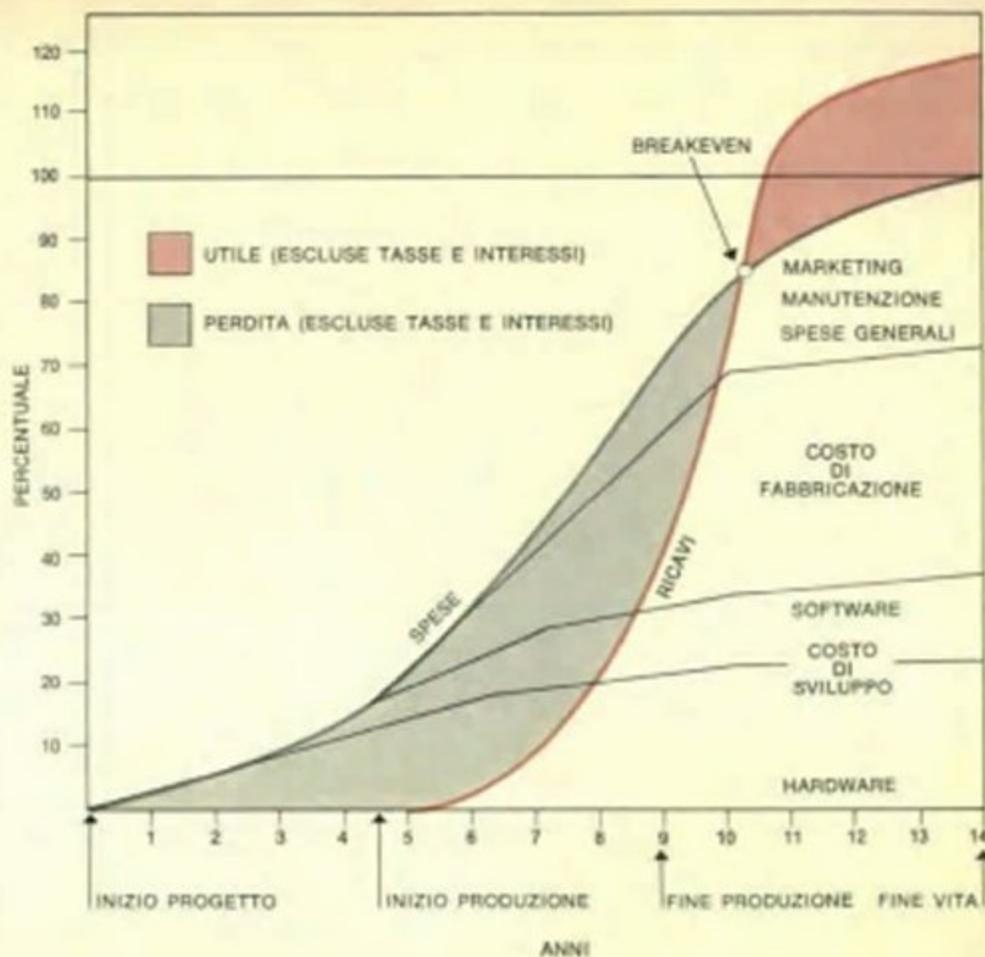
L'interazione di tipo visivo manterrà un'importanza primaria. La soluzione tecnologica corrente, mediante il tubo a raggi catodici di tipo televisivo, ha una serie di inconvenienti e limitazioni. Un progresso significativo sarà la introduzione di schermi veramente « piatti », dotati inoltre di memoria intrinseca (tali cioè che, una volta acceso un punto dello schermo, esso rimane luminoso senza bisogno di rigenerarlo periodicamente, come occorre invece fare nei normali tubi a raggi catodici). Sono allo studio diverse soluzioni in questo senso, di cui una (schermo « a plasma ») è illustrata nella figura in questa pagina.

Un altro tema importante per i terminali è la capacità di stampa. In questo campo un indirizzo preciso è la realizzazione di stampanti silenziose, in sostituzione delle fastidiose stampanti a impatto oggi in uso. Esiste una lunga lista di soluzioni per la « stampa non a impatto » basate su principi fisici e chimici i più diversi. Citiamo come esempio di sistema ingegnoso, quello detto « a getto di inchiostro », in cui goccioline elettrizzate di inchiostro vengono deflesse elettricamente sulla carta, in modo analogo agli elettroni di un tubo a raggi catodici. Per diverse ragioni, soprattutto di costo, le stampanti non a impatto non sono ancora

entrate nell'uso corrente; una soluzione tecnologica particolarmente semplice ed economica troverebbe a sua disposizione un vasto mercato.

I terminali degli anni '70 saranno caratterizzati da una elevata modularità e predisposti per una larga serie di opzioni. Più lontana appare invece la possibilità pratica di interazione fonica diretta tra l'uomo e la macchina, benché gli studi sul riconoscimento della voce siano a buon punto. (La generazione della voce è, dal punto di vista pratico, un problema più semplice; terminali « parlanti », basati sulla registrazione di frasi prefissate, sono già in uso attualmente.) D'altra parte, non sembra essere questa la cosa più necessaria; alla diffusione della elaborazione attraverso i terminali gioveranno più che certe prestazioni raffinate, una maggiore flessibilità e una sostanziale riduzione di costo.

Negli anni a venire l'elaboratore sarà sempre più lo strumento da cui dipende l'esecuzione di molte e importanti attività. Il verificarsi di guasti che ne arrestino il funzionamento avrà perciò conseguenze sempre più gravi. D'altra parte, malgrado l'introduzione di nuove tecnologie che migliorano ulteriormente la vita media dei componenti, non è possibile, dato il grande numero dei dispositivi che costituiscono un elaboratore, avere la certezza che non si verifichi alcun guasto durante il suo funzionamento. Risulta perciò essenziale trovare il modo di ridurre al minimo i tempi di ricerca e riparazione dei guasti. In questo campo c'è stata una grande evoluzione, dai primi calcolatori a oggi. Da una diagnostica affidata sostanzialmente all'abilità ed esperienza del riparatore, si è arrivati a sistemi automatici in cui l'elaboratore stesso denuncia quale delle sue parti è guasta. Nei prossimi anni oltre a una completa autodiagnostica, verrà introdotta nelle macchine anche una capacità di autoriparazione, in modo che, se anche avviene un guasto, la macchina continui a lavorare, se pure con prestazioni ridotte, fino all'arrivo del tecnico di manutenzione che ripristina lo stato normale di funzionamento. Questa capacità, che assicura la continuità del servizio per le elaborazioni più importanti e urgenti, verrà ottenuta introducendo nella macchina delle opportune ridondanze. L'uso sistematico della ridondanza ha senso pratico solo se si dispone di circuiti altamente miniaturizzati ed estremamente economici, come è il caso dei dispositivi ad alta integrazione di cui è previsto l'avvento. La LSI risulta quindi, ulteriormente, una chiave tecnologica per il progresso dell'elaboratore.



L'industria degli elaboratori elettronici presenta caratteristiche peculiari non solo per il rapido ritmo di evoluzione tecnologica, che impone ingenti spese per la ricerca e determina una vita media relativamente breve degli impianti, ma anche per la prassi di mercato vigente. In sostanza, i fabbricanti di calcolatori forniscono un servizio e gli impianti non sono di norma venduti ma dati in locazione. Questa circostanza impone al fabbricante pesanti problemi di investimenti. Tre fattori in particolare sono da considerare sotto questo profilo. Uno è l'elevato valore economico degli impianti, costituito da componenti estremamente sofisticati e quindi costosi. Un secondo fattore, peculiare di questa industria, è il « software ». Il mercato richiede non macchine ma servizi; ciò significa che occorre fornire con le apparecchiature un corredo di programmi sempre più completi ed efficienti. Un terzo fattore è la manutenzione che è, di regola, affidata allo stesso fabbricante. Tutte queste considerazioni sono, in certo modo, riassunte in questo diagramma che rappresenta l'andamento tipico spese-ricavi di un elaboratore elettronico nell'ipotesi di almeno 3000 esemplari prodotti e di una vita media di 4 anni presso l'utente. I fattori di costo sono stati divisi in tre categorie: costo di sviluppo, costo di fabbricazione, costi di marketing, di manutenzione e spese generali. Nel costo di sviluppo è stata fatta distinzione tra « hardware » e « software ». Le curve che delimitano queste divisioni, e che si riferiscono ovviamente ai valori accumulativi, sono in effetti delle spezzate (con vertici non contemporanei). I loro segmenti corrispondono alle fasi di progettazione vera e propria e di avvio della produzione, di « continuation » fino al momento della fine produzione e di « continuation » dopo la fine della produzione (in questa fase vi sono solo spese di « software »). Si vede come il recupero delle spese avvenga molto tardi e solo se l'elaboratore ha una vita sufficientemente lunga. Il momento in cui le spese eguagliano i ricavi (« breakeven ») e si comincia ad avere un profitto, varia in funzione del numero di elaboratori di prodotti. La incidenza unitaria dei vari fattori di spesa (progetto, fabbricazione, manutenzione, vendita, spese generali) diminuisce ovviamente col numero di esemplari prodotti, ma con leggi diverse tra loro. Per esempio, le spese di progetto (« hardware » e « software ») non diminuiscono proporzionalmente perché intervengono le spese per i miglioramenti e le estensioni (« continuation ») necessarie per prolungare la vita del prodotto sul mercato. Le considerazioni presentate spiegano anche la tendenza alle grandi concentrazioni industriali.

Conclusioni

Da questa pur sommaria panoramica si può rilevare come l'evoluzione dell'elaboratore non accenni a soste o rallentamenti. Anche nei prossimi anni, nuove tecnologie e nuove concezioni

verranno introdotte. Esse contribuiranno ulteriormente ad aumentare le prestazioni, a ridurre i costi, a rendere l'elaboratore sempre più accessibile all'uso; a farne, insomma, non più uno strumento per specialisti, ma un servizio a disposizione della comunità.