

LA MEMORIA DELLA MATERIA LA MATERIA DELLA MEMORIA

ANTONIO LEPSCHY, socio effettivo
ETTORE FORNASINI, socio corrispondente

Nota presentata nell'adunanza ordinaria
del 30 ottobre 2004

Premessa

Nel titolo a chiasmo di questa nota la prima parte (cioè *la memoria della materia*) si riferisce ai fenomeni fisici il cui comportamento futuro risente di quello passato e quindi, in certo senso, ne conserva memoria; la seconda parte (cioè *la materia della memoria*) allude invece a materiali (e ai fenomeni che in essi avvengono) utilizzati per “immagazzinare” informazione (in particolare dati ed istruzioni ai fini di elaborazione automatica).

A proposito di questo secondo tema si ritiene utile porre in evidenza il fatto che, in inglese (ed anche in altre lingue, spesso sul modello dell'inglese), per indicare questo processo di inserimento, conservazione e recupero di informazione si sono usate parole diverse: in particolare *memory*, *storage* e *record(ing)*. Si potrebbe rilevare che l'uso di *memory* richiama una visione antropomorfica in base alla quale, a proposito degli strumenti fisici di cui ci si avvale nell'informatica, si parla di “cervelli elettronici” e di “intelligenza artificiale”. Secondo tale impostazione, dunque, dato che – nel caso dell'uomo – si usa la parola *memory*, se ne accetta l'uso anche nel caso della macchina.

La scelta di *storage* appare invece suggerita dalla convinzione di una sostanziale differenza fra ciò che avviene nell'uomo e ciò che avviene nella macchina: pertanto se a proposito dell'uomo si parla di *memory* e se ciò che avviene nella macchina deve essere considerato sostanzialmente diverso da ciò che avviene nell'uomo, per l'elaborazione automatica sarà preferibile parlare di *storage*.

In proposito si può ricordare che il pioniere dell'informatica Charles Babbage (1791-1871) aveva chiamato *mill* la parte di elaborazione e *store* quella di memorizzazione del suo *Analytical Engine*, ade-rendo quindi al punto di vista che non "umanizzava" la macchina che egli andava progettando e cercando di realizzare e che quindi gli suggeriva di prendere a prestito termini d'uso comune relativi a strutture artificiali: *mill* è il mulino, chiamato in causa per i suoi ingranaggi (ma forse anche perché in varie lingue il verbo impiegato per indicare la sua funzione primaria di "macinare" è utilizzato, sulla base di una suggestiva analogia, anche per l'elaborazione di pensieri e concetti ed altresì di calcoli); *store* (da cui *storage*) è il negozio che, però, in questo contesto va visto non tanto come luogo di vendita quanto sotto il suo aspetto di "magazzino" e quindi di "deposito" (pur non considerandolo solo per la sua funzione passiva di conservazione, ma avendo presente anche il fatto che le merci vi vengono sistematicamente inserite e ne vengono estratte per esitarle).

Queste considerazioni suggeriscono, su un piano scherzoso, di istituire una sorta di "proporzione" (nel senso matematico della parola) nei seguenti termini:

storage sta a *electronic computer* come *memory* sta a *artificial brain*,

da intepretare nel senso che chi predilige la visione antropomorfica preferirà parlare di cervello artificiale (ed anche di intelligenza artificiale) e, quindi, di memoria, mentre chi adotta l'interpretazione opposta userà in inglese (ma anche in contesti italiani) la parola *computer* (o la tradurrà con "calcolatore" o "elaboratore") e la parola *storage* (che, eventualmente, adatterà in italiano come "immagazzinamento").

1. *La memoria come "stato" di un sistema dinamico*

1.1 *Definizioni e classificazioni riguardanti la teoria dei sistemi*

Affrontando il primo degli aspetti considerati nel titolo di questa nota, e cioè la "memoria della materia", si osserva anzitutto che, nella costruzione di modelli del mondo fisico (ma anche nella rappresentazione di processi economici e sociali), si prendono tipicamente in esame situazioni in cui una grandezza significativa del sistema che si

sta analizzando (teoricamente detta la sua "uscita" o *output*) dipende da sollecitazioni esercitate sul sistema "dall'esterno", ciascuna delle quali, nell'ambito della stessa terminologia, viene chiamata "ingresso" o *input* del sistema.

I termini *ingresso* ed *uscita* usati in questo contesto non sono dunque le "entrate" e le "uscite" di un bilancio economico (o di un bilancio materiale) ma corrispondono piuttosto a ciò che, nel linguaggio corrente, verrebbe spontaneo indicare come causa (in particolare come causa *efficiente*, nel senso dato da Aristotele a questo aggettivo) e, rispettivamente, come effetto (di tale causa).

Per comodità di presentazione di quanto segue si ritiene utile richiamare preliminarmente alcune definizioni relative a termini che sono di comune impiego nell'ambito della "teoria dei sistemi dinamici e del loro controllo" (o "automatica", che è la disciplina di cui si occupano professionalmente gli autori di questa nota).

In tale ambito un sistema si dice *statico* (o *istantaneo*) se il valore assunto dall'*uscita* in qualsiasi istante dipende solo dal valore dell'*ingresso* in quello stesso istante e non da valori assunti dall'*ingresso* in istanti diversi. Si può dunque dire che un sistema statico sia *privo di memoria* (ed in effetti in inglese si usa spesso l'espressione *memoryless*) nel senso che il suo comportamento non è influenzato da ciò che è avvenuto prima e che, quindi, delle vicende precedenti si è perduto ogni ricordo (così come, del resto, non entra in gioco alcun fattore legato a previsioni o ad obiettivi da conseguire, riguardanti il comportamento futuro).

Un sistema si dice invece *dinamico* se il valore assunto dall'*uscita* in un dato (generico) istante risente anche dei valori assunti dall'*ingresso* in istanti diversi.

Tipicamente nei sistemi dinamici *causal*, che sono quelli abitualmente considerati nella fisica, il valore dell'*uscita* in ciascun istante dipende solo da ciò che è avvenuto fino a quell'istante e, quindi, si può dire che il sistema, in qualche modo, ne serba memoria.

In tali sistemi *dinamici causal* il valore dell'*uscita* non dipende invece da ciò che avverrà nel futuro, del quale, quindi, non si ha preveggenza¹. In proposito, volendo usare ancora i termini aristotelici

¹ È però possibile farne delle previsioni basate su più o meno attendibili estrapolazioni di come l'*ingresso* è evoluto nel passato.

tradizionali, si può dunque dire che la fisica classica tiene conto di cause *efficienti* (che possiamo identificare negli *ingressi*), in un certo senso anche di cause *formali* (tale può essere infatti considerato il modello matematico della dinamica del sistema²), forse anche di cause *materiali* ma certamente non di cause *finali*³.

1.2 Modelli ingresso-uscita: funzione di memoria (e di oblio)

Come primo esempio di rappresentazione della memoria “della materia” consideriamo la vasta classe dei sistemi dinamici a tempo continuo, causali, *lineari* ed *invarianti nel tempo* per i quali il legame fra il valore dell’*uscita* $y(t)$ all’istante t e l’andamento dell’*ingresso* $u(\cdot)$ che ha agito fino all’istante t è espresso dalla formula:

$$y(t) = \int_{-\infty}^t w(t - \tau) u(\tau) d\tau. \quad (1)$$

Questa nota viene presentata in una sede interdisciplinare ed avrebbe l’ambizione di poter fornire almeno un’intuizione di quanto vien detto anche a chi, pur operando ad alto livello in campi culturali diversi, non ha familiarità con certe nozioni dell’analisi infinitesimale. Per suggerire un’interpretazione di ciò che significa la (1) si potrebbe far riferimento a fenomeni “*a tempo discreto*” anziché “*a tempo continuo*”, cioè a situazioni in cui l’*ingresso* $u(\cdot)$, l’*uscita* $y(\cdot)$ e la funzione $w(\cdot)$ sono successioni di valori che vengono assunti da dette entità in istanti discreti, numerabili come ...-2, -1, 0, 1, 2 ... (mentre non ha significato domandarsi che cosa avvenga fra l’uno e l’altro di questi istanti; si noti inoltre che la $w(\cdot)$ viene considerata solo per valori non negativi del suo argomento).

² Cfr. A. LEPSCHY, “*Ghiribizzo a proposito delle nozioni di causa formale e di analogia*”, Atti dell’Istituto Veneto di SS. LL. AA., Tomo CLXII (2003-2004), Classe di Scienze Fisiche, Matematiche e Naturali, p. 525-536.

³ È però anche vero che nelle leggi fisiche si è ritenuto di poter incorporare cause finali formulando alcune di tali leggi in termini di principi che permettano di dedurle dalla minimizzazione o massimizzazione di certe entità (e ciò, del resto, è alla base della contrarietà che i più rigidi meccanicisti hanno spesso dimostrato nei riguardi delle citate formulazioni nelle quali avvertivano aspetti metafisici che non ritenevano di poter accettare).

In tale contesto in luogo della (1) si scriverà la:

$$\begin{aligned} y(i) &= \sum_{-\infty}^i w(i-k) u(k) = \\ &= w(0)u(i) + w(1)u(i-1) + w(2)u(i-2) + \dots \end{aligned} \quad (2)$$

che non ci si illude risulti molto più chiara ma il cui significato può essere facilmente illustrato con un esempio: se la $w(\cdot)$ assume successivamente i valori 1; 1/2; 1/4; 1/8 etc., la (2) ci dice che, in un generico istante (discreto) i , l'*uscita* $y(\cdot)$ assume un valore che è la somma dell'ultimo valore dell'*ingresso* $u(\cdot)$, della metà del penultimo, di un quarto del terz'ultimo e così via; nell'istante successivo il nuovo valore dell'*uscita* avrà la stessa struttura ma ciascun valore dell'*ingresso* sarà pesato in base ai valori della $w(\cdot)$ che gli competono tenendo conto della nuova distanza di ciascuno di tali valori dell'*ingresso* dall'istante in cui si calcola l'*uscita* e quindi quel valore dell'*ingresso* che nell'istante precedente veniva moltiplicato per 1, verrà ora moltiplicato per 1/2; quello che veniva moltiplicato per 1/2 verrà ora moltiplicato per 1/4 e così via: a mano a mano che ci si sposta in avanti nel tempo, ciascun valore (precedente) dell'*ingresso* contribuirà alla formazione dell'*uscita* con un peso sempre minore.

Ritornando alla (1) – ma considerazioni analoghe valgono anche a proposito della (2) – si può dire quindi che la struttura della funzione $w(\cdot)$ è intimamente connessa con la memoria del sistema (o, se si preferisce, con la sua perdita di memoria, tanto che $w(\cdot)$ è chiamata, a volte, anche *funzione di oblio*).

Infatti nei sistemi *asintoticamente stabili*, il valore di $w(t)$ tende a 0 al tendere di t all'infinito (come in effetti avveniva per la $w(\cdot)$ considerata nell'esempio citato) rendendo così sempre meno influenti, sullo specifico valore $y(t)$ assunto dall'*uscita* all'istante t , i valori $u(\tau)$ assunti dall'*ingresso* in un istante τ che precede di molto l'istante t (cioè tale che $t-\tau$ sia grande e, quindi, per quanto si è detto, che $w(t-\tau)$ sia diventata piccola).

Da questo punto di vista si può dunque affermare che il sistema perde a mano a mano memoria di ciò che è avvenuto molto tempo prima (e la perderebbe completamente se $w(\cdot)$ fosse identicamente nulla oltre un certo valore del suo argomento).

D'altra parte, il sistema conserva memoria dell'*ingresso* che lo sollecita per un intervallo temporale di durata corrispondente a quella

dell'intervallo durante il quale $w(\cdot)$ conserva valori apprezzabilmente diversi da zero. È questo il motivo per il quale della $w(\cdot)$ si parla – come si è detto – sia come “funzione di *oblio*” (con riferimento al fatto che, oltre un certo valore del suo argomento, essa diventa trascurabile) sia come “funzione di *memoria*”, con riferimento al fatto che, sino a quel valore del suo argomento, essa si conserva apprezzabile⁴.

C'è però da notare che queste considerazioni riguardano essenzialmente la “durata” della memoria, cioè la lunghezza dell'intervallo temporale oltre il quale il sistema perde traccia della sollecitazione subita, ma non la sua “dimensione”, che si associa intuitivamente alla quantità di dati che il sistema è in grado di ritenere simultaneamente⁵.

1.3 Nozione di “stato” di un sistema in termini di classi di equivalenza di Nerode degli ingressi

Attenendoci allo schema che identifica un sistema dinamico con il legame causale fra *ingresso* ed *uscita*, in modo che il valore dell'*uscita*

⁴ Si può anche notare che tale funzione non è sempre monotonicamente decrescente e quindi si possono avere fenomeni, del resto non rari in certe patologie della memoria, in cui si ha difficoltà a ricordare un passato recente mentre si serba ricordo molto più vivo di un passato più lontano (anche se non di quello lontanissimo).

⁵ Con riferimento alla formula (1), si potrebbe osservare che una funzione $w(\cdot)$ che valesse 1 per ogni valore del suo argomento comporterebbe una memoria di durata infinita, ma di capacità molto piccola: secondo la (1) al generico istante t l'*uscita* è l'area sottesa dal grafico dell'*ingresso* fino a quell'istante e l'unico dato da “immagazzinare” all'istante t per calcolare l'*uscita* un istante dopo è il numero che esprime il valore dell'area in t . Il sistema, quindi, non è altro che un “contatore”, come quelli domestici relativi al consumo dell'acqua, del gas o dell'energia elettrica, nei quali l'unico dato memorizzato all'istante t è l'integrale dell'*ingresso* fino a quell'istante. Diverso è invece il caso di una $w(\cdot)$ che valga 1 sull'intervallo $[0,1]$ e sia nulla altrove. In esso la memoria del sistema ha durata unitaria, dato che l'*uscita* al generico istante t risente solo dell'andamento dell'*ingresso* sull'intervallo da $t-1$ a t . Tuttavia, poiché l'*uscita* all'istante t rappresenta l'area sottesa dal grafico dell'*ingresso* sull'intervallo $[t-1, t]$, l'*uscita* in un istante $t + \Delta$, prossimo e successivo a t (ovvero l'area sottesa dal grafico dell'*ingresso* sull'intervallo $[t-1+\Delta, t+\Delta]$) dipende dall'andamento di detto *ingresso* sia in prossimità di t , sia in prossimità di $t-1$. Per calcolare l'*uscita* negli istanti dell'intervallo $[t, t+1]$ si dovranno via via utilizzare gli (infiniti) valori assunti dall'*ingresso* in $[t-1, t]$ dei quali si dovrà quindi conservare memoria all'istante t . Come si vede, memoria “corta” (cioè di breve durata) non significa necessariamente memoria “piccola”.

in qualsiasi istante non sia influenzato dai valori dell'*ingresso* negli istanti successivi, la teoria astratta dei sistemi dinamici fornisce una precisa definizione della nozione di “*stato*”, che rappresenta un’interessante concettualizzazione della “*memoria*” del sistema.

Scelto un istante di riferimento t e indicando come “*passato*” (rispetto a tale istante) l’insieme degli istanti che lo precedono e come “*futuro*” l’insieme comprendente t e gli istanti successivi, la dipendenza dell’*uscita* futura dall’*ingresso* passato in generale non presenta (come si dice in termini formali) caratteri di *iniettività*, nel senso che possono esistere più andamenti dell’*ingresso* passato, pur diversi fra loro, che inducono il medesimo comportamento futuro. In altre parole all’istante t il sistema non conserva memoria *perfetta* delle sollecitazioni cui è stato precedentemente sottoposto, ma ritiene di tali sollecitazioni soltanto alcuni lineamenti essenziali.

Ciò autorizza ad aggregare in un’unica “*classe di equivalenza*”, detta “di Nerode”, tutti gli andamenti dell’*ingresso* (fino all’istante t) che danno luogo alla medesima *uscita* da t in avanti. Viene quindi chiamata “*stato del sistema*” (all’istante t) ciascuna delle classi di equivalenza nelle quali risulta partizionata la totalità degli *ingressi* potenzialmente applicabili fino all’istante t .

Gli stati forniscono così un’efficace rappresentazione della memoria, nel senso che essi racchiudono tutta e sola l’informazione concernente il passato necessaria per individuare, insieme con la sollecitazione futura, l’*uscita* futura, e la dinamica di un sistema sottende in tal modo una memoria che è tanto più “capace” quanto più grande (in un senso da precisare opportunamente) è il numero delle classi di Nerode.

Per operare concretamente con le classi di equivalenza nelle quali è partizionato un insieme, risulta quanto mai opportuno “*parametrizzarle*”, dotando di un “*nome*” ciascuna di esse, e ciò può essere ottenuto selezionando un particolare elemento entro ogni classe ed attribuendogli la funzione di “*rappresentare*” l’intera classe⁶.

⁶ Così avviene, ad esempio, con i numeri razionali, ciascuno dei quali corrisponde ad un’infinità di coppie di interi: quelle formate, appunto, da interi che stanno in uno stesso rapporto; abitualmente si sceglie come *rappresentante* di questa classe, formata da infinite coppie, una sola di esse, per esempio quella che ha il denominatore di valore assoluto più piccolo: così, ad esempio, il numero razionale corrispondente a tutti i rapporti fra interi in cui il secondo è il doppio del primo, viene abitualmente indicato con $1/2$ anziché con $2/4$ o $3/6$, e così via.

Alternativamente, si può individuare una *mappa* che associa ad ogni elemento dell'insieme sul quale è definita l'equivalenza, un ben definito oggetto di un altro insieme e che abbia la proprietà di essere un “*invariante completo*” per tale equivalenza, nel senso che essa associa il medesimo oggetto ad elementi equivalenti ed oggetti diversi ad elementi non equivalenti.

Per indicare gli *stati di Nerode* di un sistema dinamico viene comunemente adottata questa seconda alternativa, dato che spesso l'*ingresso* applicato fino all'istante t determina in tale istante un “*vettore*” di grandezze fisiche che dipendono soltanto dalla classe di equivalenza cui l'*ingresso* in questione appartiene e tale che da questo vettore è possibile risalire, senza ambiguità, alla classe stessa: per rappresentare lo stato ci si serve allora del vettore in questione.

Ovviamente, quanto più numerose sono le componenti del vettore, tanto più ricco è lo stato e più complessa la memoria del sistema.

A titolo di esempio consideriamo il moto (nello spazio tridimensionale) di una massa puntiforme sollecitata da una forza; nello studio di tale dinamica il *vettore di stato* è costituito dalle sei coordinate che forniscono, per l'istante t , le tre componenti della posizione della massa nello spazio "fisico" tridimensionale e le tre componenti della sua velocità secondo le direzioni alto/basso, destra/sinistra, avanti/indietro: tutte le possibili “storie” di sollecitazioni cui la massa è stata sottoposta prima di t e che hanno dato luogo, in tale istante, alla medesima posizione ed alla medesima velocità determinano (per ogni andamento futuro della forza applicata alla massa) la stessa evoluzione dinamica dall'istante t in poi e quindi appartengono alla medesima classe di equivalenza di Nerode. I lineamenti essenziali della storia “vissuta” dalla massa prima di t sono così riassunti nei valori di tali sei coordinate e queste rappresentano, dunque, la “*memoria*” del passato e possono venir assunte come “*stato*” del sistema ai fini di calcolarne l’evoluzione futura.

1.4 Modelli ingresso-stato-uscita

L'impostazione dello studio dei sistemi dinamici che si avvale delle classi di Nerode non è, in realtà, la più pratica dal punto di vista applicativo e ha l'inconveniente, rispetto ad altre alternative, di non

includere taluni fenomeni che, anche a livello intuitivo, si associano al dispiegarsi della memoria.

A differenza dei *modelli i/o* (cioè “*input/output*”) fin qui considerati, come sono, sostanzialmente, la (1) o la (2), l’assiomatizzazione dei “*modelli di stato*” più comunemente adottata ipotizza fin dall’inizio tre grandezze primarie (cioè *non definite con riferimento ad altre*):

- *l’ingresso*, che sollecita il sistema dall’esterno,
- *lo stato*, che viene modificato nel tempo da tale *ingresso*, e
- *l’uscita*, determinata *staticamente* (cioè solo dai valori assunti nell’istante t dalle componenti dello *stato*, e quindi in modo piuttosto banale).

In tale impostazione il legame fra *l’ingresso* e *l’uscita* non è più espresso in maniera diretta, con una formula del tipo della (1) o della (2), ma è mediato, per così dire, da una variabile interna, *lo stato*, e la struttura più importante del modello è proprio rappresentata dal legame *ingresso/stato*, che deve essere *causale*, nel senso che *lo stato*, in un generico istante t , non dipende dai valori dell’*ingresso* che sarà applicato da t (incluso) *in avanti* e deve quindi “*separare*” il passato dal futuro, cioè far sì che la conoscenza dello *stato* in t sia sufficiente per determinare *lo stato* (e quindi *l’uscita*) in ogni successivo istante t' , pur di conoscere l’andamento dell’*ingresso* da t (incluso) fino a t' (escluso).

È proprio per questo che, come si è detto, *lo stato* compendia le informazioni relative al passato utili al fine di prevedere il futuro (sotto l’azione di *ingressi futuri* che siano noti *a priori*).

Si deve però ricordare che, come pure si è detto, all’istante t il sistema può esser pervenuto ad uno *stato* particolare in infiniti modi diversi (cioè sotto l’azione di *ingressi* diversi), ma ha perduto completamente il ricordo di quale sia, fra queste sollecitazioni passate, quella che effettivamente ha avuto luogo. Si ha dunque a che fare con un tipo di memoria che perde traccia delle specifiche evoluzioni che hanno portato ad un determinato *stato* ma che serba perfettamente il ricordo di quest’ultimo.

Se così avvenisse anche nella società umana, a molti potrebbe far comodo che negli altri fosse andato perduto il ricordo di come essi hanno raggiunto determinate posizioni. Viceversa altre persone potrebbero dolersi del fatto che la società non possa ricordare più nulla degli sforzi che esse hanno dovuto compiere per conseguire tali posizioni.

È opportuno mettere in rilievo, infine, che la struttura dei modelli

li di *stato* è più complessa di quella considerata con la partizione degli ingressi nelle classi di Nerode.

In primo luogo, possono esistere *stati* del sistema non “raggiungibili”, ovvero configurazioni di memoria che non sono determinabili dalla applicazione di alcun ingresso. Queste situazioni si producono, tipicamente, quando sul sistema agiscono anche altre variabili (spesso con caratteristiche di “rumore” o di “disturbo”) che lasciano traccia o “ricordo” della loro azione.

In secondo luogo, le variabili esterne (*l'ingresso*, che supponiamo di poter manipolare a piacimento, e *l'uscita*, che supponiamo di poter rilevare in modo completo) non consentono in generale di ricostruire lo *stato*: nella memoria potrebbero infatti esistere dati “sepolti”, inaccessibili all’osservazione e che non siamo in grado di rilevare dalla sola conoscenza dell’*ingresso* e dell’*uscita*, e cioè dal comportamento “esterno” del sistema.

1.5 Memoria e previsione

Per comodità di esposizione ci si limita qui a considerare la categoria, del resto estesa ed importante, dei modelli ingresso-stato-uscita in cui la variabile di *stato* è un vettore di dimensioni finite.

L’evoluzione futura del sistema abbandonato a se stesso, cioè con *ingresso* identicamente nullo da un certo istante in poi, dipende unicamente dallo *stato* in cui il sistema si trova in quell’istante.

Sotto opportune ipotesi (in corrispondenza alle quali si parla anche di variabili di *fase* anziché di variabili di *stato*) lo *stato*, in ciascun istante, può essere espresso come l’insieme dei valori assunti in quell’istante dall’*uscita* e da un certo numero di sue derivate successive. Per esempio, nel caso semplice di un punto materiale mobile lungo una retta e descritto da una dinamica del secondo ordine, lo *stato* in un certo istante è fornito dalla posizione e dalla velocità del punto stesso in quell’istante⁷ e l’evoluzione successiva – in assenza di forze agenti dall’esterno (*ingressi*) – dipende solo da quei due valori, dei quali si

⁷ Si tratta, evidentemente, della particolarizzazione al caso unidimensionale (moto su una retta) della situazione precedentemente considerata di un punto materiale in grado di spostarsi nello spazio tridimensionale, per il quale, quindi, è necessario conoscere le tre componenti della posizione e le tre componenti della velocità.

può quindi dire che il sistema serba memoria.

Per i modelli considerati, in condizione di evoluzione *libera* (cioè con *ingressi* identicamente nulli), le equazioni del modello e la conoscenza dello *stato* raggiunto in un istante t_0 consentono di determinare tanto lo *stato* in qualsiasi istante $t_p < t_0$ che appartenga al passato di t_0 tanto quanto in un qualsiasi istante $t_f > t_0$ che appartenga, invece, al suo futuro.

Questo tipo di comportamento è conforme alla visione deterministica che si attribuisce (piuttosto arbitrariamente) a Newton e che ha avuto la sua enunciazione più esplicita e più nota con Laplace (sia pure in un contesto in cui alla supposta validità di principio dell'affermazione viene contrapposta l'irrealizzabilità pratica dell'obiettivo).

Secondo questa impostazione, la ricostruzione del passato e la previsione del futuro sono equivalenti. La prima può essere considerata una forma di *memoria* (o, almeno, un'alternativa alla memoria che ne raggiunga gli scopi con mezzi diversi); la seconda qualche cosa di sostanzialmente diverso.

Nel caso di modelli in cui lo *spazio di stato* sia completamente raggiungibile e lo *stato* sia completamente ricostruibile, se si fa invece riferimento all'evoluzione *forzata* (cioè alla presenza di un *ingresso* non identicamente nullo), fra la *ricostruzione del passato* e la *previsione del futuro* si manifesta una sostanziale differenza che può esser messa in rilievo in base alle seguenti considerazioni.

Per quanto riguarda la ricostruzione del passato, date le equazioni del sistema dinamico e noti lo *stato* in un istante t_0 e l'andamento dell'*ingresso* $u(t)$ da t_p ($< t_0$) a t_0 , lo *stato* può essere univocamente ricostruito in qualsiasi istante dell'intervallo fra t_p e t_0 (e quindi ciò può esser fatto anche per l'*uscita*). In molti contesti l'ipotesi secondo la quale, all'istante t_0 , sia noto anche l'andamento dell'*ingresso* fra t_p e t_0 appare pienamente accettabile.

Il problema di previsione su ciò che avverrà per lo *stato* dopo t_0 sarebbe analogo nel caso in cui fosse noto l'andamento futuro dell'*ingresso*, ma questa situazione di sicura conoscenza di future sollecitazioni cui sia sottoposto il sistema dopo l'istante t_0 , nella maggior parte dei casi non è soddisfatta.

Si può così concludere che il problema della *previsione* (del futuro) si pone concretamente in modo diverso da quello della *ricostruzione* (del passato). Come si è visto, però, ciò non avviene per motivi analitici o formali ma per il fatto che, in molti casi, l'andamento dell'*ingresso* può

essere noto per quel che riguarda il passato e solo in un numero molto minore di casi (e comunque con maggiore incertezza) per ciò che riguarda il futuro: tutt'al più si possono considerare, per usare la terminologia dei futurologi, diversi *scenari* in corrispondenza a ciascuno dei quali fare delle previsioni rinunciando però a sapere quale di questi si verificherà in concreto.

1.6 *Un cenno alla memoria nei sistemi “ereditari”*

Il modello *i/o* rappresentato dalla (1) o dalla (2) è *lineare* (ed *invariante nel tempo*) ma molte delle considerazioni fin qui esposte valgono anche per numerosi modelli *non lineari*. Fra questi ultimi, però, ve ne sono alcuni nei quali si manifestano fenomeni che forse assomigliano di più a certe forme di memoria degli esseri viventi.

In questi casi spesso si parla di fenomeni *ereditari* quali sono, al livello più semplice, quelli che si manifestano in presenza di *isteresi*. In questa specifica situazione la risposta del sistema nell'intorno destro dell'istante t_0 (cioè nel suo “futuro”) non è determinato solo dalle grandezze che vengono considerate come *stato* del sistema (nell'istante t_0) ma altresì dalla specifica evoluzione che ha portato a quello *stato*. In altri termini, i parametri ai quali ci siamo finora riferiti parlando di stato del sistema non sono sufficienti a “separare il passato dal futuro” (e, quindi, non costituiscono veramente uno *stato*, secondo la assiomatizzazione introdotta) e permane una sorta di sottodeterminazione dell'evoluzione futura rispetto alla specifica dei loro valori⁸.

La teoria dei sistemi ereditari è piuttosto complessa e non sembra

⁸ Un esempio semplice è quello di un sistema dinamico a retroazione in cui viene inserito un relé con isteresi, cioè un relé che commuta dall'una all'altra delle sue posizioni non in corrispondenza ad uno stesso valore (per esempio 0) del suo *ingresso* ma in corrispondenza al valore $+\Delta$ dell'*ingresso* per il passaggio dell'*uscita* dal valore inferiore a quello superiore ed al valore $-\Delta$ per il passaggio inverso. Un caso un po' più complicato, ma forse più noto, è quello dei cicli di isteresi dei materiali ferromagnetici che, ad esempio, non possono essere portati al valore nullo dell'induzione semplicemente annullando il campo magnetico, ma solo imponendo a quest'ultimo un'oscillazione smorzata in corrispondenza alla quale il punto rappresentativo delle condizioni di magnetizzazione del materiale sul piano HB descrive una sorta di spirale che potremmo interpretare come una successione di cicli di isteresi di ampiezza decrescente.

qui il caso di cercare di darne una presentazione sul piano intuitivo. Ci si limiterà quindi a mettere in evidenza che, in questi sistemi, il comportamento “futuro” risulta influenzato in maniera determinante dalla specifica evoluzione che ha portato il sistema alla situazione attuale (e non solo da tale situazione) e che quindi, in tali sistemi, si serba memoria del passato in maniera più completa di quanto avvenga per i sistemi dinamici “semplici” dei quali si è parlato in precedenza.

2 *Sistemi materiali artificiali idonei a ricevere, conservare ed estrarre informazione*

2.1 *Qualche premessa*

La tecnica, in forme sempre più evolute, ha messo a punto sistemi di inserimento, conservazione e recupero di informazioni di vario tipo a proposito dei quali si parla di registrazione (*recording*⁹) e/o di immagazzinamento (*storage*) o memoria (*memory*).

La paleoantropologia ha trovato reperti che consentono di accettare che già in tempi preistorici, prima dell’invenzione della scrittura, l’uomo ha registrato dati numerici (in un sistema di numerazione *unario*) mediante tacche su un osso o su un ramo.

Analoghe finalità di memorizzazione (anche se queste non erano forse lo scopo predominante) potrebbero essere attribuite a certe pitture preistoriche.

A rappresentazioni pittografiche si collegano le prime forme di scrittura mediante ideogrammi, nelle quali l’immagine di un oggetto o di un essere vivente stava per il suo nome, poi anche per quelli di azioni connesse a tale entità o di concetti più astratti ad essa riferibili.

Il passo successivo è stato quello della scrittura alfabetica, nella quale ciascuna “lettera” schematizza la figura di un oggetto nel cui nome il fonema iniziale corrisponde alla lettera che si vuole rappre-

⁹ La parola inglese ha lo stesso etimo della parola italiana “ricordare”, dal latino *recordari*, e si collega all’antica concezione secondo la quale l’organo della memoria sarebbe stato il cuore; del resto in francese si usa tuttora l’espressione *par cœur* dove in italiano diremmo “(imparare) a memoria”.

sentare (quasi anticipando la pratica dello *spelling* nella conversazione telefonica in cui si evitano possibili equivoci sui singoli fonemi di una parola – o, meglio, sulle lettere con i quali vengono trascritti – nominando una città sufficientemente nota di cui quella lettera sia l'iniziale, o anche usando qualche altra parola standardizzata, come ad esempio *alpha* per "a", *beta* per "b", *charlie* per "c", *dog* per "d" etc.).

In questa sezione, in cui (dopo aver trattato della memoria della materia) vogliamo parlare della "materia per la memoria", faremo riferimento al problema dei materiali "sui quali scrivere", rinunciando a trattare quello, ben più interessante ma qui "fuori tema", dei vari tipi di scrittura, in particolare di quelli alfabetici.

Nell'antichità i materiali in questione sono stati anche la pietra, il bronzo e la terracotta ma, per gli usi pratici, soprattutto l'argilla cruda o la parte interna della corteccia di certi alberi¹⁰.

2.2 *Dal papiro alla carta*

Nell'area mediterranea i tipi più importanti di materiale scrittorio adottati per testi di una certa lunghezza furono però, come è ben noto, il papiro e poi anche la pergamena, per non parlare delle tele di lino o delle tavolette cerate per note rapide.

Queste ultime costituiscono forse il primo esempio di memorie cancellabili (*erasable*) e riscrivibili (come si direbbe nel gergo dell'informatica) se si prescinde dalla pratica dei *palinsesti*, consistente nel raschiare le pergamene già scritte per riutilizzarle (spesso con una non completa cancellazione del testo iniziale ciò che, fortunatamente, in molti casi ha permesso di ricostruire tale testo).

Dall'antichità ci viene anche la distinzione fra il *volumen*, cioè il rotolo di papiro avvolgibile intorno all'una e/o all'altra di due asticciola poste agli estremi della lunga striscia, ed il *codex*, formato da pagine separate, in generale di pergamena, tenute insieme da un'opportuna rilegatura che, nelle forme più semplici, poteva consistere solo di due tavolette fra le quali i fogli erano inseriti e che venivano

¹⁰ Più precisamente quel complesso di tessuti caratterizzati dalla presenza di elementi conduttori cribrosi disposti fra la corteccia e la parte legnosa del tronco, che veniva essiccata per usarla a scopo scrittoriale e che, nella terminologia tecnica della botanica, conserva ancora il nome corrispondente al latino *liber*, poi passato ad indicare il libro.

legate con un nastro o una cordicella.

Chi si occupa delle tecniche di memorizzazione proprie dell'informatica noterà che il *volumen* può essere assimilato a quel tipo delle memorie per gli elaboratori (per esempio quelle a nastri magnetici) che bisogna far scorrere meccanicamente fino a che non compaia sotto la testina di lettura la posizione dell'informazione desiderata, così come un rotolo completamente avvolto doveva venir svolto fino alla pagina desiderata anche quando non lo si voleva leggere per intero fino a tale pagina. Invece il *codex* rilegato può essere aperto in corrispondenza ad una pagina qualsiasi, specie con l'aiuto di segnalibri, senza doverlo scorrere fino a quella pagina e quindi, per usare anche in questo caso la terminologia dell'informatica, può essere assimilato ad una memoria ad accesso casuale (RAM: *random access memory*).

Il materiale scrittorio principe della civiltà moderna, la carta, è stato inventato dai cinesi, che lo hanno ottenuto prima a partire da stracci di seta (cioè a base di fibre di origine animale) e poi, dal secondo secolo d.C., anche da stracci di tessuti formati da fibre di origine vegetale. La carta è passata poi, nel sesto secolo, ai coreani ed ai giapponesi e nell'ottavo secolo ai popoli di cultura islamica, passando infine da questi, attraverso la Sicilia e la Spagna, ai paesi dell'Europa occidentale e centrale; la prima importante cartiera dell'Europa cristiana sorse a Fabriano nel 1276.

Nel caso della carta viene spontaneo associarla alla scrittura con l'inchiosto (a mano: prima con la penna d'oca, poi con i pennini metallici ed infine con le penne a sfera; oppure con le macchine da stampa che, prescindendo dai precedenti delle tecniche xilografiche medievali, vanno dai torchi per la stampa a caratteri metallici mobili, alle *monotypes* ed alle *linotypes*; o con le macchine per scrivere o, infine, con le moderne stampanti a getto d'inchiostro); per l'esecuzione di disegni, ma anche per rapidi appunti, l'alternativa all'inchiostro è data dalla grafite e dai suoi predecessori (piombo, carboncino, sanguigna ed altri tipi di matite colorate¹¹).

¹¹ Si può ricordare, in proposito, che la parola italiana "matita" deriva da *ematite*, minerale di ferro di colore rossiccio, usato per disegnare; per tale motivo, conformemente all'uso puristico del toscano, la parola "matita" dovrebbe riferirsi solo a quelle colorate, mentre per quelle di grafite si dovrebbe usare la parola *lapis*, ripresa direttamente dal latino senza modificarne la desinenza.

2.3 Schede e bande perforate

La carta, però, è la protagonista anche di un'altra forma di registrazione di dati codificati: quella delle schede perforate (*punched cards*).

I primi inizi in questo campo riguardano l'industria tessile nel cui ambito, nel 1725, il francese Bouchon aveva inventato un sistema per telai basato sull'impiego di "aghi" e "arpini" e di una striscia continua di carta sulla quale erano praticati dei fori. Notevoli miglioramenti furono apportati al sistema fra il 1728 ed il 1734 da Falcon che, in luogo della striscia continua, utilizzò singoli cartoni (gli antenati delle vere e proprie "schede perforate") allacciati fra loro in modo da formare una catena continua che si svolgeva e riavvolgeva su due "lanterne" a quattro facce. Da questo sistema si scostò nel 1740 Régnier, che sostituì i cartoni con cilindri. Importanti miglioramenti vennero poi introdotti da Vaucanson, famoso anche come costruttore di automi, e la soluzione definitiva, in campo tessile, venne brevettata nel 1801 da Joseph-Marie Jacquard per il telaio che porta il suo nome (come del resto i tessuti che con esso vengono prodotti); tale telaio fu messo a punto da Jacquard in successive fasi fino al 1808.

Nel campo che riguarda più direttamente questa nota, le schede perforate dei telai Jacquard vennero prese in considerazione da Charles Babbage per le sue macchine che prefigurano i moderni elaboratori. Le schede perforate vennero poi inserite da Herman Hollerith nei dispositivi meccanografici da lui realizzati per il censimento del 1890 negli Stati Uniti e sono rimaste a lungo un valido strumento di memorizzazione anche dopo l'avvento degli elaboratori elettronici.

Oltre che nelle schede perforate la carta è stata largamente impiegata anche per le strisce sulle quali venivano tracciati prima i punti e le linee dell'alfabeto Morse usato per la trasmissione telegrafica e poi anche le lettere della trascrizione "in chiaro" e, soprattutto, per il vero e proprio nastro perforato (o "banda" perforata), introdotto nel 1846 da Alexander Bain ed adottato nel 1857 da Wheatston, sempre in campo telegrafico. Tali nastri di carta hanno poi trovato largo impiego anche nelle macchine meccanografiche, nelle telescriventi e negli apparati telex.

Oggi però, in campo informatico, la carta rimane confinata all'eventuale registrazione finale "in chiaro" di un documento (quando

non ci si accontenti di leggerlo sullo schermo del *computer*) mentre ha perduto qualsiasi funzione nel trattamento dell'informazione, che avviene su altro tipo di supporto all'interno dell'elaboratore.

Analogamente sono scomparsi dall'uso dispositivi di memorizzazione meccanica, anche di tipo digitale ma soprattutto di tipo analogico, come le *cammes* degli automi di Vaucanson o dei Jacquet-Droz o, in tempi più recenti, le sagome (*templates*) dei torni automatici a copiare.

2.4 *Registrazione di immagini e di suoni*

Durante il secolo diciannovesimo la tecnica ha portato alla ribalta due innovazioni fondamentali nel campo della *memorizzazione* o, più precisamente, della registrazione (*recording*): quella di immagini con tecniche fotografiche e quella del suono con tecniche fonografiche.

A proposito della fotografia si può dire che essa affronta in modo nuovo (riducendo al minimo l'intervento dell'operatore umano, limitato alla scelta dell'inquadratura e, al più, del tempo di posa e dell'incisione, nonché alla messa a fuoco) un problema già tradizionalmente risolto *ab antiquo* con la pittura, il disegno e con le "arti grafiche" che consentivano di ottenere numerose repliche di uno stesso originale, come nel caso della xilografia, dell'incisione o dell'acquaforte.

Si deve però osservare che le tecniche fotografiche dell'Ottocento erano decisamente innovative per quanto riguarda l'aspetto chimico ma si avvalevano di strumenti, come la *camera obscura* ed i sistemi di lenti, già ben noti da un lato ai fisici e dall'altro ai pittori ed ai disegnatori.

Per quanto riguarda i fenomeni chimici, agli alchimisti era già noto l'effetto della luce sul cloruro d'argento; un altro materiale con caratteristiche analoghe di fotosensibilità era il cosiddetto bitume di Giudea.

Il primo ad avviare ricerche in campo fotografico fu J. Nicéphore Nièpce, che ottenne i principali risultati fra il 1816 ed il 1822; un ulteriore determinante contributo fu conseguito nel 1835 da J. Daguerre, che introdusse le tecniche di sviluppo in grado di rendere stabile l'immagine (e sul cui nome fu coniato il neologismo la cui forma italiana è dagherrotipo). Altri contributi significativi si debbono a W. H. Talbot, all'astronomo Herschel (che, fra l'altro, ha coniato il termine *fotografia*), ad H. Bayard, ad Abel Nièpce (nipote di Nicéphore), che introdusse l'impiego di lastre di vetro, ad F.S. Larcher, a

G. Le Gray, ad R.L. Maddox ed a G. Eastman, fondatore della Kodak, cui si deve l'impiego di pellicole in rullo, su supporto in celluloide.

Alla fine del secolo diciannovesimo, dopo vari tentativi ed esperienze, si sarebbe arrivati, con il cinematografo dei fratelli Lumière, alla memorizzazione e riproduzione di immagini in movimento (sfruttando il fenomeno della persistenza retinica e quindi utilizzando una successione in rapida cadenza di immagini fotografiche fisse).

Ben diversa era la situazione per la riproduzione del suono che, verso la fine dell'Ottocento, attuò per la prima volta il sogno, fino a quel momento considerato irrealizzabile, di conservare e riprodurre segnali acustici, in particolare la voce umana (per la quale non era possibile avvalersi di sistemi come quelli dei *carillons* o degli organetti meccanici, che del resto non riproducevano un suono preesistente ma generavano, con un loro timbro particolare, un motivo inciso su di essi con opportune tecniche).

In particolare nel 1877 Thomas Alva Edison realizzò il fonografo a rullo e Charles Cros (brillante figura di medico, poeta, attore ed inventore) presentò all'*Académie des Sciences* di Parigi il principio del disco fonografico, che venne poi attuato con il grammofono brevettato nel 1887 da Emil Berliner.

La registrazione e riproduzione di immagini in movimento sincronizzate con quella di suoni darà anche luogo, nel 1927, al cinema sonoro nel quale la cosiddetta colonna sonora, registrata direttamente sulla pellicola, veniva affiancata alle immagini garantendo così il sincronismo, praticamente quasi impossibile da conseguire quando il suono da associare all'immagine è memorizzato su un supporto diverso da quello usato per l'immagine stessa (per esempio su un disco fonografico).

Per la registrazione e riproduzione dei suoni, accanto ai classici dischi (di vari materiali ed infine, tipicamente, di vinile¹²) che rap-

¹² La matrice per i dischi di Berliner era di zinco: la registrazione era ottenuta sovrapponendo al disco uno strato di cera che veniva inciso dalla puntina; attaccando il disco con un'acido, questo scioglieva lo zinco dove la cera era stata asportata dalla puntina; la matrice così ottenuta consentiva poi la stampa delle copie, che in un primo tempo erano di gomma lacca e poi vennero realizzate in acetato (eventualmente su supporto di vetro o di metallo). Risale al 1947 l'impiego del vinile per i dischi, di grande diametro, detti LP (*long playing*) o "a 33 giri" (precisamente a 33+1/3 giri al minuto, mentre lo standard precedente era a 78 giri); i dischi a 45 giri (di diametro minore) risalgono al 1949.

presentarono a lungo la soluzione più diffusa, si può ricordare che nel 1898 il fisico danese Valdemar Poulsen aveva proposto la registrazione di un segnale acustico sotto forma di magnetizzazione di un idoneo supporto materiale, che però all'epoca poteva essere realizzato solo in forma di fili metallici (tipicamente di ferro o sue leghe), in generale poco pratici. Un deciso miglioramento si ottenne con la registrazione su nastro magnetico; il brevetto di Fritz Pleumer è del 1928 ma importanti contributi furono dati, dopo vari tentativi, soprattutto da Eduard Schiller e Friedrich Matthias nel 1934.

Dopo l'occupazione della Germania al termine della seconda guerra mondiale tecnici militari americani vennero a conoscenza dei progressi conseguiti in quel paese nel campo della registrazione magnetica e ne diffusero la conoscenza negli Stati Uniti dove si dedicarono alla ricerca, allo sviluppo ed alla produzione su larga scala l'Ampex e la 3M. Nel 1954 vennero immesse sul mercato bobine di nastro magnetico ed apparecchiature per la loro registrazione e per l'ascolto. Nel 1963 la Philips mise in commercio la prima *musicassetta*. Nel 1975 la JVC realizzò il sistema VHS (*Video-Home System*) per registrazioni *audio* e *video* su nastri *magnetici* (e contemporaneamente venivano messe in commercio le cosiddette *videocassette* già registrate).

Di altre tecniche più recenti per la realizzazione di analoghi obiettivi si dirà più avanti.

2.5 Le soluzioni tecniche a mano a mano utilizzate per la memorizzazione in campo informatico

Il nastro magnetizzabile ebbe larghissimo impiego, oltre che per la registrazione di suoni, anche nel campo dell'informatica, in particolare, come si è già ricordato, per le memorie di massa.

Non si trattò, tuttavia, della prima delle soluzioni adottate in questo campo, che ora si passeranno brevemente in rassegna considerando i veri e propri elaboratori, a partire dalle esperienze pionieristiche avviate già negli anni Trenta del Novecento, e dalla realizzazione dei primi calcolatori elettronici, degli anni Quaranta e dell'inizio degli anni Cinquanta, per i quali furono escogitati vari tipi di memorie, alcuni dei quali oggi ci appaiono assai curiosi.

Nel primo modello di elaboratore digitale di Konrad Zuse, lo Z1,

concepito già nel 1936, si utilizzavano memorie meccaniche nelle quali, per ciascun *bit*, venivano utilizzate tre piccole piastre metalliche forate da spostare opportunamente, realizzando una memoria di 64 word di 22 bit ciascuna.

Nei cosiddetti *relay computers* di R. Stibitz, in particolare nel suo *K Model*, si utilizzava un relé per ogni bit. Poco dopo, per memorie di lavoro di modesta capacità, si utilizzarono anche insiemi di flip-flop (realizzati a tubi elettronici; più tardi anche a transistor).

Fra le soluzioni che oggi ci possono sembrare più curiose si possono citare le *linee di ritardo* acustiche (impiegate da Maurice Wilkes, dell'Università di Cambridge per l'EDSAC, ultimato nel 1949). Su queste linee veniva immessa una sequenza di impulsi acustici che, in uscita, erano poi trasdotti (mediante un dispositivo piezoelettrico) in impulsi elettrici, successivamente rigenerati eliminando le distorsioni, amplificati ed infine ritrasdotti (mediante un dispositivo piezoelettrico funzionante in modo simmetrico rispetto al precedente) in impulsi acustici riportati all'ingresso della linea di ritardo. In tal modo il segnale continuava a circolare indefinitamente nella linea e la specifica informazione di interesse veniva "letta" prelevando gli impulsi corrispondenti ad essa in base al loro indirizzo temporale. Per la linea di ritardo si utilizzarono tubi di vetro, di adeguata lunghezza, riempiti di mercurio. Accanto agli effetti piezoelettrici si sfruttarono anche fenomeni di magnetostrizione (su fili metallici).

In un altro tipo di queste memorie (oggi obsolete) piccole superfici metalliche disposte sullo schermo di un tubo a raggi catodici (CRT), funzionando come armature di condensatori, venivano caricate (o meno) dal fascio elettronico modulato che spazzolava lo schermo. La lettura risultava "distruttiva" (in quanto, per leggerne il contenuto, ciascun condensatore doveva venir scaricato; era quindi indispensabile rigenerare continuamente l'informazione memorizzata). Queste memorie a tubi catodici vennero adottate nel 1946/47 da F.C. Williams e T. Kilburn. L'indirizzo intrinseco di questo sistema di memorizzazione era spaziale bidimensionale (consisteva infatti nelle coordinate del singolo microcondensatore sullo schermo del CRT) ma veniva trasformato in una coordinata temporale (unidimensionale) dal moto di spazzolamento del pennello elettronico.

La soluzione più classica per le memorie negli elaboratori della prima generazione fu però il tamburo magnetico, inventato nel 1932 da

G. Taushek ma impiegato per la prima volta in calcolatori, ancora a livello sperimentale, già nel 1948 ma solo nel 1953 in un prodotto commerciale, all'epoca di avanguardia, come il 650 dell'IBM: esso constava di un cilindro rotante ad elevata velocità, rivestito di un *film* magnetico sul quale una fila di testine di lettura e scrittura, in posizione fissa rispetto al cilindro rotante, esplorava lo stato di magnetizzazione di ciascuna delle minuscole areole della superficie lungo una circonferenza del tamburo (lettura) o lo modificava, portandolo al valore voluto (scrittura).

Come nel campo della memorizzazione di suoni si era passati dal cilindro di Edison al disco di Berliner (per altro già preconizzato da Cros), così anche in campo informatico si passò da tamburi a dischi magnetici (e poi anche ottici) di varie dimensioni e per diverse finalità; già negli anni Cinquanta del Novecento, nel RAMAC dell'IBM (1956) si combinarono in una pila cilindrica numerosi dischi rotanti sui quali l'indirizzo di ciascuna informazione era fornito da un coordinata relativa al disco che contiene tale informazione, da un'altra coordinata spaziale relativa alla circonferenza volta a volta di interesse su tale disco (e che veniva fornita ad una testina spostabile in senso radiale), ed infine da una coordinata temporale, relativa all'istante in cui sotto la testina passa la parte della traccia contenente l'informazione di interesse.

Con gli elaboratori di seconda generazione si passò poi, nel 1951, a memorie veloci ad accesso casuale costituite da *matrici* di anelli di ferrite, con caratteristica di magnetizzazione a ciclo di isteresi praticamente rettangolare, secondo il suggerimento di W.J. Forrester che sarebbe poi diventato più noto per i suoi studi di *System Dynamics*.

Il passo successivo corrisponde all'utilizzazione di memorie elettroniche a semiconduttori (realizzate inizialmente dalla INTEL) passate man mano a livelli di integrazione sempre più spinti (ed oggi elevatissimi).

2.6 Alcune considerazioni ulteriori a proposito delle memorie per l'informatica

In un moderno sistema informatico si manifestano esigenze differenti per i diversi tipi di memorie, cosa della quale, del resto, ci si è resi conto da molto tempo ed in qualche caso fin dai primordi delle realizzazioni in questo campo.

Si può pensare anzitutto alla distinzione fra le memorie per immagazzinare istruzioni e quelle per immagazzinare dati, che erano

tenute distinte nella cosiddetta “architettura Harvard” ma che von Neumann genialmente suggerì di mantenere unificate¹³. Resta comunque il fatto che, per scopi differenti, si dispone di memorie di tipo diverso e che una distinzione fondamentale è quella fra le memorie ad accesso diretto (alla singola informazione), cioè le RAM (distinguibili in *statiche* e *dinamiche*, queste ultime essendo quelle che richiedono tecniche di continuo “rinfresco” (*refresh*) dei dati memorizzati), e le memorie ad accesso seriale.

In base ad un criterio distinto si individuano le ROM (*read only memory*) o memorie a sola lettura, registrate una volta per tutte (e quindi adatte per certi insiemi fondamentali di istruzioni) alle quali si affiancano memorie che, pur non prestandosi ad una continua riscrittura, possono consentire, con idonei procedimenti, la scrittura a cura dell’utente (PROM: memorie programmabili a sola lettura), o anche la cancellazione (EPROM: *erasable PROM*, sia con procedimenti più laboriosi, sia con tecniche di più facile impiego come le EEPROM: *electrically erasable PROM*).

All’interno dei vari tipi di elaboratori c’è poi tutta una gerarchia delle memorie, che vanno dai *registri* della memoria di lavoro, realizzati con RAM statiche, alla eventuale memoria *cache*, alla memoria principale, realizzata con RAM dinamiche, alla memoria secondaria (generalmente a dischi magnetici) fino alle eventuali memorie *off-line*. Passando dalle prime alle ultime i costi (per bit memorizzato) descrescono ma i tempi di accesso (e le dimensioni di ingombro) aumentano¹⁴ cosicché l’architettura della memoria va

¹³ Si potrebbe però osservare che in alcuni elaboratori per scopi particolari (*special purpose*) come per esempio nei PLC (*controllori a logica programmabile*) risulta tuttora conveniente adottare l’architettura Harvard.

¹⁴ In proposito può essere interessante ricordare che, nell’imaginario collettivo, anche a livello di vignettistica, gli elaboratori elettronici sono stati spesso rappresentati emblematicamente dalle unità nastri che, ai tempi del loro maggiore impiego, avevano capacità di immagazzinamento superiori a quelle di qualsiasi altro dispositivo di memorizzazione e ricorrevano a tecniche sofisticate, lavorando praticamente sotto vuoto, per garantire la costanza della velocità lineare di passaggio del nastro sotto la testina (cioè che comportava velocità angolari variabili per le bobine dalle quali il nastro si svolgeva e sulle quali, rispettivamente, si riavvolgeva). Queste unità, fra l’altro, assumevano un aspetto vagamente antropomorfizzabile, dato che le due bobine di ciascuna di esse, che apparivano attraverso una finestratura, facevano pensare ai due occhi di un androide.

opportunamente studiata in modo da sfruttare al meglio le caratteristiche di ognuno dei tipi indicati ai fini dei compiti che essi devono svolgere.

2.7 *Dischi magnetici ed ottici*

Non è facile, in fatto di memorie, sviluppare un discorso sistematico ed organico, in quanto si possono seguire criteri diversi, per esempio in base al significato di ciò che si memorizza: immagini, suoni, dati, istruzioni; oppure in base agli scopi: di intrattenimento, di documentazione, di elaborazione; o, ancora, in base alle tecnologie o ad altri criteri. In certi contesti i criteri indicati risultano accoppiabili in un certo modo; in contesti diversi gli stessi accoppiamenti possono risultare privi di significato o addirittura falsanti.

In questa nota nella sezione 2.4 si è parlato della registrazione di suoni e di immagini e nelle sezioni 2.5 e 2.6 (da due distinti punti di vista) delle memorie utilizzate in campo informatico. Ora si parlerà di dispositivi di memoria con caratteristiche (almeno parzialmente) diverse da quelle finora utilizzate, che però hanno trovato impiego (sia pure con qualche differenza, più però di tipo pratico che di principio) tanto nella registrazione di suoni ed immagini prevalentemente a scopo di intrattenimento, quanto in quella di istruzioni e dati in campo informatico.

Si tratta di "dischi" ad elevata densità di immagazzinamento, di tipo magnetico e di tipo ottico; nei primi è legata allo stato di magnetizzazione assunto da ciascuna areola di una pellicola di materiale magnetizzabile che ricopre il disco; nei secondi dalla pervietà ottica (cioè dalla presenza o meno di forellini) delle corrispondenti areole. Lo stato di magnetizzazione è rilevato da apposite testine di lettura; nei dischi ottici da elementi fotosensibili (per esempio fotodiodi o fototransistor) raggiunti o meno da un raggio di luce (per esempio laser).

In linea di principio i dischi magnetici sono sostanzialmente analoghi ai tamburi ed ai dischi RAMAC di cui si è già parlato; in pratica, dal punto di vista tecnico e da quello delle prestazioni, se ne differenziano in misura importante. Basti considerare che il sistema RAMAC è del 1956 e che i cosiddetti *floppy disc* risalgono al 1978,

quando vennero commercializzati dalla Apple e dalla Tandy. Poco dopo furono introdotti gli *hard disc*.

La tecnologia del disco ottico è stata sviluppata dalla Philips a partire dalla fine degli anni Sessanta del Novecento, ma la sua concreta diffusione commerciale è più recente. L'impiego di questa tecnologia per registrazioni di suoni ed immagini in movimento (cioè per la realizzazione del cosiddetto videodisco) risale, a livello sperimentale, al 1972 ma la sua diffusione commerciale data dal 1978.

I CD (*Compact Disc - Digital Audio System*) sono il risultato di una collaborazione fra la Philips e la Sony e sono stati commercializzati sia nella forma ROM sia in quella RW (cancellabili e riscrivibili). Con la sigla CDI ci si riferisce a quelli interattivi.

La sigla DVD corrisponde alla denominazione *Digital Versatile Disc*; questi sono disponibili dal 1997 in diverse varianti (ROM, R (registrabili), WO (*write-once*), ed altre, fra le quali i DVD interattivi).

2.8 *Un rapido cenno alle tecniche di memorizzazione più recenti ed alle prospettive future*

Attualmente vengono considerate interessanti le prospettive di sviluppo delle memorie a bolle magnetiche, di quelle criogeniche e di quelle basate sull'effetto Josephson.

Altrettanto si può dire di una sorta di resurrezione delle schede perforate, per altro con una profonda trasformazione di dimensioni e di prestazioni, quale è costituita dal *Millipede* della IBM, una scheda su supporto polimerico di dimensioni ridottissime.

Sembra più appropriato parlare invece di tempi più lunghi (ed anche di una maggiore incertezza sia nella previsione di questi tempi, sia in quella delle specifiche caratteristiche e prestazioni di questi dispositivi) a proposito delle memorie olografiche, di quelle dette "a risoluzione atomica", di quelle basate sull'impiego di colle epossidiche ed infine di quelle a base proteica. Con queste ultime si potrà forse dire che l'uso del termine "antropomorfizzante" *memoria* non è più basato su una analogia con ciò che avviene nel cervello dell'uomo (o, più in generale negli animali) ma su una corrispondenza più stretta.

Conclusioni

La parola italiana ‘*memoria*’, usata due volte nel titolo di questa nota, così come quelle delle altre lingue moderne che hanno la stessa etimologia, è legata a parole del latino e del greco che si riferivano essenzialmente ad attività della mente umana.

Nel campo di quell’area scientifica e applicativa che oggi è ormai abituale denominare “scienza e tecnologia dell’informazione” la parola ‘memoria’ è stata largamente impiegata, in sensi che abbiamo cercato di chiarire e con riferimento a contesti anche abbastanza diversi fra loro, assieme ad espressioni più o meno equivalenti, quali in particolare ‘immagazzinamento (delle informazioni)’, ‘registrazione’ e altre analoghe.

In questa nota si sono analizzati due distinti campi di uso della parola ‘memoria’ nell’area dell’informazione e precisamente quello che riguarda la teoria dei sistemi dinamici e del loro controllo (e cioè l’*automatica*) e quello che riguarda l’elaborazione dell’informazione (e cioè l’*informatica*).

Come si è messo in evidenza, questi due usi sono abbastanza diversi.

Nel primo campo ci si riferisce ad effetti di sollecitazioni che si propagano nel tempo e che quindi si manifestano ancora dopo che le sollecitazioni che li hanno provocati sono ormai cessate, talché si può dire che, in qualche senso, di tali sollecitazioni si conservi il ricordo (anche se, come si è visto, si tratta di una memoria un po’ particolare, che non distingue fra le sollecitazioni che appartengono alla stessa classe di equivalenza).

Nel campo dell’informatica ci si riferisce, invece, alla possibilità di “scrivere”, “conservare” e “leggere” simboli associati ad informazioni, il che avviene alterando lo stato fisico di un supporto materiale, con l’eventuale avvertenza di rigenerare periodicamente questo stato, in caso esso si degradi, e rilevandolo in modo da ottenere un segnale idoneo a fornire l’informazione che era stata immagazzinata.

In questo campo, nell’ultimo mezzo secolo, si sono compiuti progressi enormi in particolare in fatto di densità di immagazzinamento (bit per centimetro quadrato) e di velocità di scrittura e lettura, ciò che ha avuto conseguenze molto notevoli anche sulla vita di tutti i giorni, non solo per ciò che riguarda le applicazioni vere e proprie del-

l'informatica ma anche le attività di intrattenimento (come può esser reso evidente confrontando le voluminose "pizze" delle pellicole cinematografiche che venivano portate nelle sale pubbliche di proiezione ed i piccoli DVD che ci consentono di vedere gli stessi film sul televisore domestico).

RIASSUNTO

La nota analizza, in due parti distinte, il funzionamento di sistemi dinamici che risentono, nella loro evoluzione, di sollecitazioni ricevute nel passato (ed a proposito delle quali si può quindi dire che ne serbino memoria) e le caratteristiche e proprietà dei principali materiali e dispositivi utilizzati dall'antichità più remota sino ad oggi per registrare informazioni di vario tipo, con particolare (ma non esclusivo) riferimento al loro impiego nel campo dell'informatica, per la memorizzazione di dati e di istruzioni.

ABSTRACT

The first part of the paper is devoted to the analysis of dynamical systems performance. Specifically, in these systems the effects of past inputs on the present state of the system may be regarded as a form of memory (whereas an instantaneous system can be considered as memoryless).

The second part of the paper describes the standard storage devices adopted in the field of computer technology as well as other storing devices adopted for recording words, sounds and images.