LA MACCHINA DELLA MENTE

Parte I^{a}. Il cervello e il calcolatore.

Autore: Renato Nobili
Dipartimento di Fisica "G.Galilei" – Università di Padova.

1.0. Sommario: In questa parte sono trattate alcune questioni fondamentali per la ricerca sull'intelligenza artificiale (IA). Dopo una breve analisi storico-critica del paradigma logico booleano, sono portati allo scoperto gli aspetti teorici più interessanti che riguardano il funzionamento della macchina mentale con particolare riguardo ai processi autoriflessivi che caratterizzano il cervello umano come macchina introspettiva. Sono infine introdotti gli argomenti che inducono ad abbandonare i tentativi di risolvere i problemi dell'IA nell'ambito di una teoria dei processi seriali a favore di una nuova teoria dei processi paralleli.

1.1. Il sogno di Boole

Per quasi tutto il quarantennio successivo all'apparizione dell'articolo *A Logical Calculus of Ideas Immanent in Nervous Activity* di McCulloch e Pitts (1943) il neurone è stato considerato come un *decisore a soglia*, ossia un dispositivo che genera in uscita un segnale a gradino quando una somma algebrica di segnali positivi o negativi applicati in ingresso supera un certo valore. In quel periodo la complessità dei processi nervosi e le possibilità dell'elettronica erano nettamente sottovalutate. Il ruolo funzionale del neurone poteva apparire tanto elementare quanto quello di una valvola termoionica o, dopo la sua scoperta avvenuta all'inizio degli anni 50, di un transistor. Poiché s'ignorava che anche il tessuto di sostegno e alimentazione della rete nervosa (la glia) scambia segnali con i neuroni, l'informazione nervosa sembrava essere unicamente processata dalle cellule nervose ed era naturale ritenere che i processi mentali fossero riconducibili all'interazione di unità funzionali simili ai componenti di un circuito elettronico (Arbib, 1965). L'analogia sembrava corretta anche perché si riteneva che i segnali nervosi si propagassero elettricamente attraverso i contatti sinaptici intesi come punti di saldatura tra le cellule nervose.

La localizzazione della memoria appariva invece problematica: alcuni ritenevano che l'informazione nervosa fosse memorizzata nei valori di soglia dei neuroni, altri che fosse codificata in molecole contenute nei neuroni, magari nello stesso DNA. Delle diverse ipotesi allora formulate rimane valida ancora oggi quella di Hebb (1949), che fu il primo ad ipotizzare che la memoria sia dovuta alla formazione o all'alterazione dei contatti sinaptici che mettono in comunicazione le cellule di un popolazione neuronale.

Per capire la differenza tra la visione di allora e quella dei nostri giorni basta pensare che oggi la complessità funzionale di un singolo neurone è paragonata da alcuni autori a quella di un microprocessore. In realtà il paragone è sbagliato perché il neurone ha un comportamento probabilistico mentre il microprocessore è deterministico.

L'aspetto più suggestivo di quella visione tanto ottimistica era fornito dal modello della risposta neuronale di tipo *tutto o niente*. Questo semplice principio di funzionamento trovava legittimazione teorica nella rappresentazione in cifre binarie (*bit*), ossia mediante successioni di 0 e 1, dell'informazione intesa come una sequenza di decisioni dicotomiche (Shannon, 1949). Tutto ciò corrispondeva pienamente a quella che poteva sembrare la profetica eredità culturale di George Boole (1853), fondatore del calcolo logico: l'idea che le regole di formazione del linguaggio logico (affermazione, negazione, congiunzione, disgiunzione e loro combinazioni), in definitiva le stesse operazioni fondamentali del pensiero, fossero equivalenti a un algoritmo algebrico operante su variabili proposizionali a due valori: *tutto* = 1 = *vero*; *nulla* = 0 = *falso*.

In effetti, se si assume che i costituenti elementari delle reti nervose funzionino come decisori a soglia, è naturale ipotizzare che nel cervello avvengano processi analoghi a quelli del calcolo di

Boole. Si può immaginare quale sia stato il divertimento di McCulloch e Pitts, e dopo di loro di Kleene (1956) che riprese la questione in modo più completo e rigoroso, nel riuscire a dimostrare che tutte le operazioni del calcolo booleano possono essere effettuate da reti formate da decisori a soglia temporalmente ritardati (De Luca e Ricciardi, 1981).

D'altronde, poiché il funzionamento di un decisore a soglia può essere approssimato tanto quanto si vuole da un circuito formato da operatori booleani elementari e dispositivi di ritardo temporale, si deduce che ogni elaborazione d'informazione effettuabile mediante decisori a soglia può essere simulata da una rete sufficientemente complessa del secondo tipo. Insomma, se le reti neurali di McCulloch e Pitts fossero davvero buoni modelli delle reti nervose reali, i cervelli animali e i calcolatori digitali fornirebbero due modi perfettamente equivalenti di processare l'informazione; con tutte le conseguenze, anche sul piano filosofico, che si possono immaginare. L'impetuoso sviluppo delle tecnologie informatiche degli anni '60 contribuì ad alimentare ulteriormente quest'illusione di semplicità.

Di fatto, mediante processi booleani ricorsivi applicati a sequenze digitali fu possibile implementare il calcolo aritmetico e quindi calcoli matematici d'ogni genere e complessità. La questione del rapporto tra *logica* e *aritmetica* si collegò in modo naturale alla stupenda problematica relativa alla ricorsività algoritmica: l'analisi di Gödel, le macchine di Turing, la teoria dei linguaggi formali, ecc. Sembrò così potersi effettuare anche la "quadratura" del cervello: che nulla potesse accadere nella mente che non potesse essere simulato da un calcolatore sufficientemente complesso.

L'indagine sui fondamenti del pensiero sembrava chiudersi così con un circolo perfetto. Si poteva pretendere di più? Quali dubbi avrebbero potuto sorgere circa la validità di tale approccio dal momento che matematici come von Neumann, Shannon, Moore, Arbib, ecc. contribuirono a legittimarlo elevandolo al rango di una teoria sistematica e completa? Tale concezione ha informato fino a tempi recenti la ricerca sull'*intelligenza artificiale* (IA), a proposito della quale merita ricordare i nomi di Minsky, Newell, Simon.

Purtroppo i risultati in questo campo sono stati tanto deludenti quanto le prospettive sembravano promettenti. Detta brutalmente, dopo 30 anni di intense ricerche il software prodotto dagli studiosi dell'IA non ha ancora fornito qualcosa che assomigli vagamente a un'implementazione del pensiero umano.

1.2. Gödel, Turing e von Neumann

Alcuni logici e teorici dell'IA hanno nutrito per alcuni decenni un ambizioso progetto: spiegare come un processo di elaborazione dell'informazione possa diventare capace di "autoriflessione", di riprodurre, cioè, la capacità del pensiero umano di pensare il proprio pensato, in altri termini di generare l'*autocoscienza*. Quest'idea è stata divulgata in forma avvincente e suggestiva da uno di loro, Douglas Hofstadter, nel libro *Gödel, Escher, Bach* (1979). L'autore trae ispirazione dalle proprietà *autoreferenziali* dell'aritmetica scoperte da Kurt Gödel (1931). Ma anche, sebbene l'autore non lo dichiari esplicitamente, da quella, del tutto affine, dell'*autoriproducibilità* di automi dotati di capacità costruttive, che fu delineata da John von Neumann (1966) in una serie di conferenze e appunti tra il 1949 e il 1956, anno, questo, della sua morte.

Nelle visioni di Gödel e di von Neumann le possibilità, rispettivamente, dell'autoreferenza e dell'autoriproduzione dipendono dalla condizione che i sistemi considerati - rispettivamente, l'algoritmo aritmetico e l'automa costruttore - siano sufficientemente complessi. La questione è così importante da meritare un'ampia digressione.

Per non spaventare il lettore poco familiare con la matematica eviterò di entrare nei dettagli tecnici della questione. I concetti che cercherò di presentare meritano una certa attenzione se non altro perché segnano la linea di demarcazione tra la scienza del presente e quella del passato. Per scandire meglio la sequenza logica del discorso userò gli ordinali latini e scriverò in corsivo i termini che veicolano le definizioni importanti.

I. L'aritmetica e l'algebra elementare sono due *algoritmi* molto popolari. Anche le regole per la costruzione di figure geometriche mediante riga e compasso costituiscono un algoritmo, sebbene sia basato su operazioni di disegno invece che su formalismi alfabetici e numerici. Con la massima generalità possiamo dire che un algoritmo è un insieme di regole procedurali che si applicano, in modo generalmente *ricorsivo*, ai dati di un certo insieme per produrre risultati che vanno ricollocati nello stesso insieme. Un processo ricorsivo può arrestarsi ad un certo punto quando il risultato desiderato è stato ottenuto, o continuare idealmente per sempre al fine di produrre un risultato sempre meglio approssimato. Nella pratica matematica gli algoritmi sono elaborazioni simboliche che permettono di produrre enti matematici, geometrici e fisici utili alla scienza.

II. Ora diciamo che un algoritmo ne *interpreta* un altro se, operando su un proprio insieme di dati o simboli, corrispondenti all'insieme di dati o simboli dell'altro, esso è capace di produrre risultati esattamente corrispondenti a quelli dell'altro. Brevemente, se è in grado di *simulare* i comportamenti dell'altro. Un algoritmo che non sia abbastanza complesso riesce, al più, ad interpretare algoritmi simili o più semplici. Per esempio, le operazioni di addizione e sottrazione riescono a interpretare le moltiplicazioni e le divisioni mediante l'uso di esponenziali e logaritmi.

Un algoritmo sufficientemente complesso può essere *universale*, vale a dire capace di interpretare qualsiasi altro algoritmo. È stato merito di Alan Turing (1936) l'avere dimostrato che una macchina dotata di una semplice testina mobile capace di leggere e scrivere alcuni simboli su un nastro di lunghezza infinita, memorizzando temporaneamente pochissimi dati, è in grado di simulare qualsiasi altra macchina calcolatrice di complessità e memoria interna arbitrariamente grandi. In particolare, *la macchina di Turing universale può eseguire ogni sorta di calcolo aritmetico*. Il prezzo da pagare è solo il tempo di calcolo. Su questa base si dimostra anche che qualunque macchina di Turing capace di eseguire calcoli aritmetici è universale.

Si può comprendere quanta importanza abbia avuto per la logica matematica sapere che *l'aritmetica è un algoritmo universale*. È questa la ragione per cui tutte le procedure matematiche, per esempio anche quelle puramente geometriche e persino quelle più raffinate che impegnano le menti dei più geniali fisici teorici, purché opportunamente interpretate, sono riconducibili, in ultima analisi, a puri e semplici calcoli aritmetici.

III. D'altronde ogni algoritmo possiede una struttura che può essere presentata in forma assiomatica ed elaborata con procedimenti logici. Precisamente come un sistema puramente formale costituito da *nozioni primitive, assiomi* e *regole d'inferenza*. Per esempio, l'algoritmo geometrico, che fa materialmente uso di riga e compasso, ha una struttura che si compendia in forma assiomatica nella geometria euclidea, intesa come sistema logico finalizzato alla dimostrazione di verità geometriche.

Componendo le nozioni primitive secondo i principi della logica formale si ottengono espressioni più o meno complesse. Tutte le espressioni che si possono ottenere per combinazione logica di altre, indipendentemente dalla verità di queste altre, si dicono *formule ben formate* (nel caso della geometria si potrebbe dire *figure ben disegnate*).

Formule complesse d'uso frequente possono essere opportunamente rimpiazzate, per definizione, da formule (o nozioni) più semplici. Gli assiomi sono formule ben formate assunte *vere* per definizione. Ad esempio: "ogni numero ha un successore", "per due punti passa una e una sola retta" ecc. Le formule che si possono ottenere componendo, elaborando e riducendo logicamente le formule costruibili con gli assiomi e le regole d'inferenza del sistema assiomatico considerato si chiamano *teoremi*.

In pratica la logica formalizza il ragionamento matematico. Essa stabilisce una successione di passaggi alcuni dei quali hanno l'effetto di aumentare la lunghezza delle proposizioni logiche (nel caso della geometria la complessità della figura), altri di diminuirla, fino a produrre, come risultato finale, una proposizione (una figura) che è giudicata interessante.

Le diminuzioni di lunghezza possono ottenersi applicando le regole d'inferenza del sistema assiomatico - le quali in pratica stabiliscono che certe formule possono essere rimpiazzate da altre meno complesse - oppure applicando le regole fondamentali della stessa logica, ad esempio il *modus ponens* , che permette di rimpiazzare la formula *"se A allora B"* semplicemente con "*B"* (nel caso della geometria è lecito cancellare le parti del disegno che sono già state usate per produrre un certo risultato).

L'assiomatizzazione dell'aritmetica, effettuata da Peano e Dedekind nella seconda metà dell'800, si basa sulle nozioni primitive di *numero* e *successore* e sull'enunciazione, nei termini di queste, degli assiomi e delle regole d'inferenza che istituiscono la possibilità di eseguire quante si vogliano addizioni e moltiplicazioni, persino infinite operazioni, e *inoltre* il diritto di enunciare una verità generale mediante una concatenazione infinita di verità particolari (*assioma d'induzione*). La formalizzazione logica dell'algoritmo aritmetico, esteso ai numeri reali, non è altro che l'algebra elementare (il principio d'induzione applicato alla geometria è implicito nel metodo di esaustione di Archimede).

IV. Dunque l'algoritmo aritmetico può essere considerato da due punti di vista diversi: *a*) come qualcosa di concreto, operazionale, comportamentale, diacronico; *b*) come qualcosa di astratto, rappresentazionale, strutturale, sincronico. Nel primo caso esso si presenta come un linguaggio che ha come universo del discorso *i numeri interi particolari*. Possiamo definirlo *linguaggio aritmetico*. Nel secondo caso come un linguaggio che ha come universo del discorso le *proprietà dei numeri interi*, o, in altri termini, *le classi di numeri interi*, giacché la caratterizzazione di una proprietà equivale alla definizione di una classe. Possiamo definirlo *linguaggio meta-aritmetico*. In modo simile ogni altro algoritmo si presenta, da un lato, come un sistema di procedure atte a trasformare concretamente un insieme finito di dati particolari in un altro insieme finito di dati particolari; dall'altro come una struttura formale astratta che definisce ed elabora una collezione infinita di proprietà generali.

La distinzione dell'aspetto *algoritmico* da quello *logico* caratterizza in modo irriducibilmente dualistico l'intera matematica. Storicamente, tra i due aspetti si è stabilita una drammatica tensione che si è manifestata, prima, nel tentativo, tentativo perseguito principalmente da Gottlob Frege (1879-1903), di ridurre l'aritmetica alla logica, poi nell'indagine di Kurt Gödel (1930, 1931) volta alla ricerca delle difficoltà logiche che s'incontrano nel descrivere le proprietà dell'algoritmo aritmetico. Il confronto dei due punti di vista ha avuto l'effetto di ampliare oltre i limiti delle loro formulazioni originarie prima la logica (attraverso i lavori di Whithead, Russel, Zermelo, Fraenkel, Hilbert, Tarsky ecc.) e poi l'algoritmica (con i contributi di Church, Turing, Kleene, Moore, ecc.).

V. L'indagine di Gödel si basa sulla scoperta che le dimostrazioni logiche possono essere tradotte in calcoli aritmetici. L'idea nasce dal fatto che le elaborazioni logiche concrete, ad esempio le dimostrazioni dei teoremi di un sistema assiomatico, non sono altro, in ultima analisi, che produzioni algoritmiche di dati a partire da altri dati, esattamente come lo sono i calcoli aritmetici. In altri termini, le operazioni della logica booleana, di cui consiste una dimostrazione, non sono altro che una sorta di aritmetica applicata all'elaborazione di dati consistenti di proposizioni ben formate. Ora, l'aritmetica, giacché è un algoritmo universale, può distinguere ed enumerare ordinatamente tutti gli assiomi e le proposizioni ben formate di una teoria logica. Può anche rappresentare le deduzioni dei teoremi dagli assiomi di una qualsiasi teoria logica (ad esempio la geometria euclidea o la meccanica razionale), o da altri teoremi della stessa teoria, nella forma di operazioni aritmetiche condotte sui numeri che codificano gli assiomi o questi altri teoremi. La danza delle complessificazioni e delle semplificazioni che intervengono in una dimostrazione logica si tradurrà, in corrispondenza, in una danza di allungamenti e raccorciamenti di espressioni numeriche. Procedendo in questo modo l'intero sistema dei ragionamenti logici ammissibili dalla teoria così codificata sarà rappresentato da un sistema di operazioni aritmetiche.

Il passaggio chiave del procedimento gödeliano consiste in questo: *se si applica la codificazione aritmetica alla stessa teoria assiomatica dell'aritmetica si genera un circuito autoreferenziale dotato di straordinarie implicazioni e conseguenze*.

L'algoritmo aritmetico è messo in condizioni d'interpretare se stesso e di dimostrare, nella forma di particolari calcoli aritmetici, le proprietà logicamente dimostrabili dell'aritmetica stessa. Così ad esempio il teorema "esistono infiniti numeri primi", proposizione che, si badi bene, non è verificabile analizzando caso per caso quali numeri siano primi, sarà codificata da un numero particolare. La dimostrazione del teorema si presenterà allora nella forma di una successione di passaggi aritmetici che mettono in relazione tale numero con quelli che codificano gli assiomi. Insomma in questo modo l'aritmetica, intesa come teoria logica, riesce a codificare in termini di calcoli aritmetici *particolari* le proprietà *generali* del calcolo aritmetico. Questo può essere fatto per tutto ciò che è dimostrabile nell'ambito della teoria aritmetica. Mirando a una maggiore generalità, possiamo anche affermare che mediante la *gödelizzazione* il *linguaggio* aritmetico riesce a inglobare tutte le proposizioni *logicamente dimostrabili* del suo *metalinguaggio*.

VI. La fase culminante dell'indagine gödeliana si presenta quando ci si rende conto che un metalinguaggio contiene necessariamente più proposizioni del suo linguaggio oggetto, allo stesso modo come il numero dei sottoinsiemi di un insieme contiene più elementi dell'insieme stesso. Questo fatto può essere formulato con maggiore precisione affermando che la *cardinalità* delle proposizioni del metalinguaggio è superiore a quella del linguaggio.

Due insiemi possiedono la stessa cardinalità se possono essere mappati uno sull'altro. Nel caso degli insiemi finiti la cardinalità equivale al numero di elementi. Ma la nozione di cardinalià si estende anche agli insiemi infiniti. Un insieme infinito può essere sempre mappato in una sua parte propria, tanto che un insieme può definirsi infinito se gode di questa proprietà. Ad esempio tutti i numeri interi possono essere mappati nei soli numeri pari mediante una semplice moltiplicazione per 2. Si può anche dimostrare, per esempio, che tutti i punti di un quadrato possono essere mappati su uno dei suoi lati e viceversa. Ma in molti altri casi queste mappe tra insiemi infiniti non possono farsi. Così, ad esempio, come ha insegnato Cantor, mentre i numeri interi possono essere mappati sui razionali e viceversa, non si può costruire una corrispondenza biunivoca tra i punti di un segmento e l'insieme dei numeri naturali giacché la cardinalità del continuo è superiore a quella del discreto.

Ora, il linguaggio aritmetico ha la cardinalità del discreto, perché la totalità dei calcoli aritmetici possibili può essere ordinata ed enumerata. Ma il suo metalinguaggio, avendo come universo del discorso le classi dei calcoli aritmetici possibili, ha la cardinalità del continuo. Stando così le cose, si comprende immediatamente che devono esistere infinite proposizioni del metalinguaggio che non sono interpretate dal linguaggio. Bisogna anzi aggiungere che si tratta di un'infinità infinitamente preponderante! Dato che la costruzione gödeliana mappa la parte logicamente dimostrabile del metalinguaggio aritmetico nel linguaggio aritmetico, si deduce che il metalinguaggio aritmetico possiede infinite *proposizioni indecidibili*, o teoremi indimostrabili. Il grande logico austriaco ha concretamente mostrato come si possa costruire una proposizione indecidibile.

VII. Originariamente l'analisi di Gödel riguardava il rapporto tra la struttura logica e la potenza algoritmica della matematica. Ciò rifletteva in modo naturale l'interesse maturato nel dibattito scientifico del primo '900. I risultati di Gödel furono reinterpretati da Turing come problema del rapporto tra la struttura e il comportamento di una macchina calcolatrice. La macchina di Turing riduce all'osso la sostanza di questo problema. Essa consiste di una testina che può spostarsi di un passo in avanti o di uno all'indietro su un nastro infinitamente lungo diviso in caselle uguali; oppure leggere o scrivere su queste caselle un simbolo preso da un alfabeto finito di simboli; il tutto secondo regole definite da una tabella di programmazione.

La macchina di Turing, opportunamente integrata con un dispositivo ausiliario capace di transitare tra alcuni stati e di governare il comportamento della testina in funzione di tali stati, diventa *universale*. Ciò significa che, una volta che i dati, gli stati e le istruzioni di programma di un'altra qualsiasi macchina calcolatrice siano stati codificati in un'opportuna sequenza di simboli del suo nastro, essa potrà simulare quest'altra macchina calcolatrice.

La macchina di Turing universale, benché abbia una struttura, tutto sommato, abbastanza semplice, può fare tutto quello che può essere fatto dai calcolatori più complessi. Questa possibilità dipende in definitiva dal fatto che la struttura della macchina più complessa viene codificata come un particolare insieme di dati della macchina più semplice. Grazie a questa riduzione, l'intera teoria dei processi di calcolo si riconduce allo studio delle possibilità computazionali di questo congegno di complessità matematica descrivibile in poche pagine. In pratica, l'universalità della macchina di Turing si identifica con la sua capacità di eseguire qualunque calcolo aritmetico esattamente come l'universalità dell'algoritmo aritmetico si identifica nella possibilità computazionali della macchina di Turing universale. L'analogo del teorema di indecidibilità di Gödel si ha nel fatto che una macchina di Turing universale è sempre in grado di simulare un'altra, ma, in generale, non di prevedere se un processo di calcolo effettuato da quest'altra macchina si arresterà dopo un numero finito di passi (*halting problem*).

VIII. La teoria degli automi di von Neumann ricalca lo stesso schema concettuale. Per certi aspetti, essa rappresenta una naturale estensione del punto di vista di Turing. Un *automa costruttore* è una macchina capace di usare oggetti reperibili nell'ambiente per produrre altri oggetti da riporre nel medesimo ambiente. A tale scopo esso deve disporre di un elenco d'istruzioni e possedere un repertorio sufficientemente ricco di sensori e strumenti di lavoro per eseguire tutte le necessarie operazioni di cernita e assemblaggio sulla base di dette istruzioni; perciò, in generale, deve anche essere capace di elaborare informazione. Un automa di complessità strutturale insufficiente non riuscirà a produrre altro che oggetti di complessità inferiore alla sua. Ma se l'automa è dotato di organi capaci di svolgere un numero sufficiente di funzioni elementari, ed è inoltre sufficientemente bene organizzato, allora esso diviene *universale*. Vale a dire che, leggendo un opportuno programma d'istruzioni, esso diventa capace di produrre qualsiasi oggetto, anche automi di complessità uguale o superiore alla sua; dunque anche una copia di se stesso. A questo punto è chiaro che la cellula vivente ha tutti i requisiti di un automa universale!

Von Neumann - richiamandosi al teorema di indecidibilità di Gödel - afferma che l'automa non potrà "auto-copiarsi" semplicemente esaminando direttamente la propria struttura interna mediante sensori o altri ipotetici apparati d'introspezione, poiché nessuna automisurazione o auto-osservazione può essere completa. Di fatto, nessun apparato di misura può compiere una misura dentro se stesso. Affinché l'autoriproduzione abbia luogo è essenziale che l'automa costruttore disponga di una descrizione della sua struttura interna e di un apparato di decodificazione che gli permetta di dedurre le procedure costruttive o, equivalentemente, di un elenco completo delle istruzioni riguardanti le modalità di costruzione e assemblaggio delle sue parti. Tale è appunto il codice genetico degli organismi viventi.

Si ha in ciò l'analogo della codificazione di un sistema assiomatico nei dati di un algoritmo universale, senza la quale l'interpretazione autoreferenziale non potrebbe effettuarsi. È notevole il fatto che von Neumann giunse ad affermare tutto questo circa tre anni prima della scoperta del DNA (Watson e Crick, 1953), suggerendo, tra l'altro, sulla base di argomenti relativi a condizioni di funzionalità ottimale, che l'informazione genetica sia codificata in sequenze lineari di dati (come i simboli sui nastri delle macchine di Turing). Il sommo matematico ha fornito l'esempio di un semplice automa capace di autoriprodursi costituito di cellette quadrate uguali, ciascuna dotata di 29 stati interni, tra i quali si trova a transitare interagendo deterministicamente con le cellette contigue.

Il punto cruciale di questa profonda visione, che può ritenersi l'istanza fondante della biologia teorica, è il ruolo svolto dalla distinzione tra la *complessità strutturale* e la *complessità comportamentale* di un sistema formale o naturale. Sotto un certo livello di complessità strutturale, i comportamenti di un algoritmo, di una macchina calcolatrice o di un automa costruttore sono perfettamente prevedibili e descrivibili in termini finiti sulla base dalla conoscenza della sua struttura. In queste stesse circostanze, la potenza interpretativa dell'algoritmo, quella di calcolo del calcolatore o la capacità produttiva dell'automa non può che essere limitata. Corrispondentemente, l'universalità algoritmica, quella computazionale e quella costruttiva, quindi la gödelizzazione e l'autoriproduzione, sono impossibili.

Ma se tali sistemi possiedono strutture sufficientemente ordinate e articolate si crea un fatto sconvolgente: la varietà dei comportamenti qualitativamente diversi esplode esponenzialmente verso il limite ideale del continuo, e la maggioranza di essi *non sono prevedibili e descrivibili sulla base dalla conoscenza della struttura*, per quanto questa possa essere perfetta. L'unico modo di conoscerli è quello di osservarli nel loro effettivo svolgimento. Proprio in queste stesse circostanze, la potenza interpretativa dell'algoritmo, la capacità di simulazione del calcolatore o quella produttiva dell'automa diventano universali.

1.3. Il problema dell'autocoscienza

L'idea che la capacità introspettiva del pensiero ricalchi il procedimento gödeliano è ventilata nel celebre libro di Hofstadter in un modo piuttosto oscuro. Dipanandola dalle metafore, dai doppi sensi, dai giochi di parole e dalla festosa allegria da *cartoons* che avviluppano lo spirito del libro, essa può essere brevemente descritta nel seguente modo: un cervello pensante è la sede di processi di produzione di certi stati di eccitazione della rete nervosa, che l'autore definisce *simboli attivi*. Questi sono così definiti perché, a differenza dei simboli passivi che intervengono nei formalismi matematici, non ricevono i loro significati da relazioni di corrispondenza poste da un soggetto esterno, ma sono essi stessi capaci di "attivare" e reclutare relazioni con oggetti esterni. Per mantenere l'analogia coi casi precedenti bisogna assumere che la "macchina mentale" sia capace di trasformare i simboli attivi in altri simboli attivi, secondo procedure che vengono istruite in qualche modo non conosciuto. Naturalmente, non deve esistere alcun apparato supervisore, nessun *homunculus* cosciente nascosto che legga e interpreti un elenco di istruzioni al fine di far eseguire al cervello tali operazioni: l'attività di produzione e riproduzione simbolica di cui consiste ciò che si usa chiamare *mente* deve generarsi e sostenersi autonomamente per interazione causale dei simboli attivi; simboli che si attivano e si disattivano nella struttura cerebrale secondo una specifica dinamica. In analogia coi i concetti illustrati nei paragrafi precedenti, aggiungiamo la seguente definizione: la mente A è capace di *interpretare* la mente B se A è capace di stabilire una corrispondenza tra ogni produzione simbolica di B e una delle sue.

Affinché ciò possa avere luogo, A avrà bisogno di mezzi atti ad acquisire e decodificare i simboli e le produzioni simboliche di B; dovrà pertanto possedere una capacità di rappresentazione interna coordinata ad apparati d'interazione col mondo esterno. Una mente che abbia una complessità strutturale insufficiente, ad esempio quella di una scimmia, non può interpretare altro che produzioni simboliche di complessità limitata, ma se la sua struttura è abbastanza complessa ed articolata, e la sua memoria è abbastanza capace, essa diventa una mente *universale*. Purché sia opportunamente istruita, essa potrà interpretare ogni altro genere di produzione simbolica: quella di una macchina calcolatrice, di un animale, di altre menti dotate di qualsiasi grado di complessità. E' evidente che, in questo senso, *la mente umana è universale*!

Hofstadter non spiega come in un cervello animale possa formarsi una mente universale, giacché nel libro citato, o in un altri del medesimo autore, la questione dell'universalità risulta implicita solamente nell'idea della corrispondenza coi teoremi di Gödel. Tuttavia, a parte considerazioni di carattere filogenetico, possiamo ragionevolmente ipotizzare che tale capacità si organizzi fin dalle primissime fasi dell'età evolutiva attraverso la comunicazione espressiva degli stati emotivi, l'esercizio della loro simulazione, l'apprendimento emulativo e l'assimilazione dei codici di comportamento. Ma, naturalmente, lo sviluppo di queste facoltà non basta a produrre una capacità d'interpretazione universale.

Nemmeno riguardo la genesi dell'autocoscienza, lo studioso americano fornisce spiegazioni adeguate all'idea che egli avanza. Possiamo tentare di completare in questa sede la sua visione spingendo fino in fondo l'analogia con la costruzione gödeliana e incorporando in questa i concetti chiave della teoria degli automi di von Neumann. La questione sembra riguardare il rapporto tra due fondamentali facoltà della mente umana: l'*immaginazione* (identificabile con l'attività di produzione simbolica descritta sopra) e il *linguaggio naturale*. Sembra ragionevole, infatti, stabilire una corrispondenza di queste facoltà coi termini del dualismo matematico descritto nel paragrafo precedente: l'attività algoritmica da una parte e il suo linguaggio logico dall'altro. La gödelizzazione del rapporto tra l'immaginazione e il linguaggio naturale potrebbe compiersi nel seguente modo.

L'attività pensante, che dal punto di vista strutturale sembra consistere di un sistema d'interazioni tra aree e centri cerebrali capaci di reclutare ed evocare associativamente tracce della memoria sotto forma d'immagini complesse coerenti, dal punto di vista comportamentale si presenta come produzione dinamica di fantasie specifiche particolari. Indipendentemente dalle relazioni che le immagini evocate hanno con l'esperienza percettiva originaria, un'elaborazione mentale di una fantasia ha caratteristiche formali simili a quelle di una produzione algoritmica di dati numerici. Essa è suscettibile di essere interpretata e compresa da un'altra mente sufficientemente complessa tramite una corrispondenza con le fantasie specifiche particolari di questa seconda mente. Naturalmente questo può avvenire se le due menti possiedono un apparato per la manifestazione simbolica delle proprie fantasie e un sistema di codificazione comune, dunque un linguaggio comune. Nessuna mente è in grado di percepire direttamente le fantasie di un'altra mente, ma può solo ricostruirsele nella forma di fantasie sue se i messaggi ricevuti sono sufficientemente coerenti e dettagliati.

Una mente universale potrà interpretare e comprendere, ossia esemplificare nella propria immaginazione, in virtù della sua capacità di comunicare, anche le espressioni simboliche da lei stessa liberamente o casualmente generate, indipendentemente dal fatto che le espressioni simboliche derivino da fantasie effettivamente prodottesi nella sua precedente esperienza. In altri termini, l'immaginazione e la comunicazione devono ritenersi due attività diverse e indipendenti, che trovano le loro connessioni nei processi interpretativi del pensiero. Nel loro insieme, esse impartiscono all'attività mentale un carattere dualistico simile a quello che si ritrova nella distinzione tra l'algoritmo aritmetico e la teoria aritmetica. L'autocoscienza, nella sua essenza, dovrebbe consistere proprio nella gödelizzazione dell'attività pensante, precisamente come un procedimento autoriflessivo che dovrebbe esplicarsi come interpretazione immaginativa delle elaborazioni linguistiche indotte dall'attività immaginativa stessa.

Perseverando nell'analogia, e rifacendosi alle tesi di von Neumann, dovremmo ipotizzare che una mente universale non possa giungere all'autocoscienza semplicemente "percependo" o "guardando dentro se stessa", poiché l'autopercezione diretta deve ritenersi tanto impossibile quanto l'autointerpretazione diretta della teoria aritmetica senza il passaggio attraverso la codificazione numerica delle sue proposizioni, o l'autoriproduzione degli organismi viventi senza la codificazione dell'informazione genetica in un DNA.

Ma come si ha nel caso delle macchine di Turing e di von Neumann, l'universalità interpretativa dipende dalla possibilità di rappresentare in un proprio insieme di dati i programmi che permettono di simulare altre macchine. Allo stesso modo dobbiamo assumere che l'universalità di una mente dipenda dalla possibilità di rappresentare gli stereotipi comportamentali di altre menti, come in una sorta di teatro delle marionette interno. Come una macchina di von Neumann dotata di universalità costruttiva è per questa stessa ragione capace di autoriprodursi, così una mente dotata di universalità interpretativa è capace di "autointerpretarsi". Per fare ciò essa deve utilizzare proprio quelle risorse che le permettono di interpretare le altre menti.

Dovremmo perciò ipotizzare che in una mente universale esista uno speciale sistema simbolico capace di rappresentare ed esemplificare immaginativamente anche la struttura formale della propria attività pensante, le relazioni tra le parti di questa struttura e le funzioni dei suoi meccanismi interni. Generalizzando la nozione di "simbolo del sé" introdotta da Hofstadter, possiamo chiamarlo *sistema simbolico del sé*.

In un certo senso, quest'ipotesi non fa altro che ripresentare in altra forma la tesi dell'homunculus annidato nel cervello. Solo che ora questo strano personaggio non deve essere immaginato come un secondo soggetto pensante collocato in una parte speciale del nostro cervello. Se così fosse non si potrebbe sfuggire alla regressione all'infinito che ci obbligherebbe ad ipotizzare un secondo homunculus dentro il cervello del primo, un terzo dentro la cervello del secondo ecc. Questo personaggio non deve nemmeno essere identificato con l'homunculus somato-sensoriale, che si usa dipingere sulle mappe della corteccia cerebrale per indicare dove risiedono i neuroni che corrispondono alle diverse parti del corpo. Le strutture somato-sensoriali sono un aggregato di aree specializzate costantemente impegnate a codificare e veicolare i flussi d'informazione reciprocamente indipendenti tra cervello e soma. L'homunculus di cui si ha bisogno deve essere inteso come un'entità simbolica complessa che rappresenta in modo integrato lo schema di ciò che accade nel cervello, non i processi mentali effettivi nella loro complessità. Allo stesso modo, gli assiomi dell'aritmetica teorica non rappresentano tutto ciò che l'algoritmo aritmetico può generare, e la doppia elica del DNA non rappresenta ciò che accade nell'organismo vivente.

Hofstadter non spiega dove risieda e come si generi il codice mentale della marionetta che ci rappresenta internamente. Ma se il parallelo coi casi matematici vale fino in fondo, dobbiamo assumere che nel processo di formazione dell'autocoscienza, assieme alle capacità di produzione simbolica e di comunicazione con altri esseri pensanti, si stabiliscano anche relazioni di corrispondenza proiettiva e correlazione simbolica tra i modi di comportamento e comunicazione propri e quelli altrui. Secondo questa visione, l'autocoscienza s'innescherebbe a cominciare dall'istante in cui una mente capace di generazione grammaticale e produzione simbolica, una volta che abbia raggiunto un certo livello d'organizzazione interna, diventa prima un interprete universale delle attività mentali altrui e poi un interprete autoreferenziale della propria attività mentale.

Il grande sogno dell'IA, la creazione dell'autocoscienza artificiale, sembra dunque condizionato dalla possibilità di costruire macchine elaboratrici di informazione capaci di comunicare linguisticamente con gli esseri umani, interpretarne il pensiero, introiettarne lo schema strutturale, fino ad elaborare un proprio sistema simbolico del sé e a divenire in questo modo anche capaci d'autointerpretazione. Può la moderna tecnologia dei calcolatori arrivare a tanto?

1.4. La macchina introspettiva

La capacità autoreferenziale dell'aritmetica e quella autoriproduttiva del costruttore universale suggeriscono la possibilità che una macchina di Turing universale possa essere resa autoreferenziale, vale a dire posta in condizioni di analizzare, descrivere e comunicare i calcoli che essa ha effettuato o sta effettuando e le procedure che ha usato o sta usando. Una macchina calcolatrice dotata di questa capacità potrebbe essere giustamente definita *introspettiva*. Essa fornirebbe un modello della capacità autoriflessiva del pensiero umano e una sua eventuale implementazione equivarrebbe alla creazione del pensiero artificiale. Negli anni '80 vari ricercatori hanno avviato alcune ricerche teoriche volte allo sviluppo di una teoria delle macchine introspettive (Bartlett, 1992).

In un articolo del 1985 Kurt Konolige definisce la *macchina introspettiva* come un sistema dotato di un *sottosistema di credenze*. Secondo la sua definizione, un sottosistema di credenze è una struttura computazionale all'interno di un agente artificiale capace di rappresentare un numero finito di credenze riguardanti i fatti del mondo. Questo sottosistema è in grado di accettare un quesito posto da un agente esterno, e tentare di verificare se la risposta al quesito può essere derivata dal suo sistema di credenze. Il funzionamento della macchina si basa sulla scomposizione di un quesito complesso in quesiti più semplici e sul successivo tentativo di rispondere ai sottoquesiti.

Konolige assume che la macchina introspettiva possieda per ogni sua credenza *C*, che egli definisce *non doxastica*, anche la credenza "*credo che C sia una mia credenza*", che egli definisce *doxastica*. Questo semplicemente significa che il linguaggio doxastico è il metalinguaggio di quello non doxastico. L'analogia coi casi precedentemente considerati è palese. Chiaramente, una condizione necessaria affinché la macchina sia effettivamente capace d'introspezione è che l'algoritmo che elabora i quesiti non doxastici sia universale. Il teorema d'indecidibilità di Gödel assicura allora che esistono infiniti quesiti ai quali la macchina non sa rispondere.

Konolige non lo dice, ma un'altra condizione perché la macchina possieda effettivamente la capacità introspettiva è che essa sia capace di porre autonomamente a se stessa tutti i quesiti che possono essere posti da un agente esterno. Può una macchina di Turing essere programmata per fare tutto questo? La risposta è negativa. La capacità di generare autonomamente i quesiti che possono essere imprevedibilmente posti da agenti esterni contrasta col determinismo della macchina di Turing. Questa capacità potrebbe essere supplita da un dispositivo capace di porre quesiti casuali "ben formati" sulla base di un'opportuna grammatica generativa. La difficoltà sta nel fatto che nessuna macchina deterministica è in grado di produrre una sequenza casuale di simboli.

La domanda deve dunque essere riformulata nel seguente modo: può una macchina di Turing universale equipaggiata con una sorgente di quesiti casuali essere programmata per funzionare come una macchina introspettiva? *In teoria sì*, perché la macchina di Turing universale possiede tutte le condizioni per la sua gödelizzazione. *In pratica no*, perché i tempi di calcolo dei procedimenti ricorsivi attraverso i quali essa può tentare di rispondere ai quesiti, costruendo e analizzando classi di credenze, crescono in generale esponenzialmente col numero delle credenze. Questo dipende dal fatto che la cardinalità delle classi di credenze è superiore a quella delle credenze. Il collo di bottiglia che impedisce nella pratica il processo introspettivo sta nel fatto che le macchine di Turing, e con esse ogni calcolatore di tipo ordinario, sono *sequenziali*, nel senso che esse sono costruite in modo da eseguire un numero relativamente piccolo di operazioni a ogni passo del loro orologio interno. La coscienza di un evento mentale ci servirebbe assai poco, e forse sarebbe un intralcio, se invece di giungere con un ritardo di circa mezzo secondo dopo che un'intenzione è stata elaborata inconsapevolmente dal cervello, come gli esperimenti di Libet e collaboratori (1979) hanno dimostrato, impiegasse un tempo maggiore. La sua formazione richiede che la macchina cerebrale sia molto potente, oltre che ben organizzata, per generare in tempi utili i processi introspettivi. Forse per questa ragione essa è comparsa tardi nell'evoluzione delle specie.

1.5. Processi seriali e paralleli

Limitatezza numerica delle operazioni eseguibili nell'unità di tempo e serialità dei processi di calcolo sono proprietà caratteristiche dei calcolatori ordinari e della macchina di Turing, che li rappresenta tutti in modo paradigmatico. Ciò significa che: (1) il tempo impiegato da queste macchine per effettuare un insieme di calcoli indipendenti non può essere minore della somma dei tempi necessari per effettuarli singolarmente; (2) in pratica sono possibili solo elaborazioni che richiedono tempi di calcolo che crescono come una potenza piccola del numero di dati iniziali.

Un moderno calcolatore da tavolo può eseguire alcuni miliardi di operazioni per secondo in corrispondenza dei cambiamenti di stato della sua CPU (*central processing unit*). La velocità dei calcolatori dipende dal fatto che i segnali elettrici si propagano a velocità prossime a quelle della luce. La velocità dei segnali nervosi è dello stesso ordine di grandezza di quella del suono, essendo limitata dalla natura elettrochimica dei processi da cui dipende. Questa lentezza, a fronte del vantaggio evolutivo degli organismi più veloci, ha imposto ai sistemi nervosi degli animali di adeguarsi ad un principio di *massima parallelizzazione possibile a parità di funzioni*. Nel seguire questa via le specie animali sembrano non aver perso nulla delle possibilità offerte a priori dalla natura: al fine della lotta per l'esistenza, i processi di calcolo parallelo effettuati dai sistemi nervosi si rivelano assai più efficienti di quelli che potrebbero essere eseguiti, in modo più preciso ma più lento, da macchine calcolatrici seriali.

Molti neurofisiologi concordano nell'affermare che le effettive unità di processamento dell'informazione nervosa nella corteccia cerebrale non sono i singoli neuroni, ma piccole popolazioni neuronali organizzate in colonne che contano alcune migliaia di neuroni. Dato che il cervello umano contiene circa diecimila miliardi di neuroni (10^{10}), si può stimare che esso ospiti alcuni miliardi di queste piccole PU (*processing unit*) neuronali. Lavorando alla frequenza di un centinaio di cambiamenti di stato per secondo – tali sono le frequenze massime delle oscillazioni di potenziale che si misurano nelle colonne corticali - il cervello sembra capace di effettuare, al più, alcune centinaia di miliardi d'operazioni diverse per secondo. Il cervello umano sarebbe dunque solo un centinaio di volte più potente di un moderno computer da tavolo.

Se le cose stessero semplicemente in questi termini dovremmo aspettarci il raggiungimento della parità di potenza di calcolo dei personal computer con quella del cervello umano nel giro di pochi anni. In realtà, se si mettono a confronto le capacità di memoria dei due sistemi la stima va a netto vantaggio del cervello animale. Considerando che i neuroni interagiscono in modo probabilistico con un notevole grado di dispersione statistica, non sembra sbagliato ritenere che la sinapsi di un neurone abbia una capacità effettiva di memoria dell'ordine di grandezza di un bit. Dato che un neurone ospita circa diecimila sinapsi, ne consegue che la memoria umana ha una capacità di circa

100 terabyte (1 byte = 8 bit; 1 mega = 10^6, 1 giga = 10^9, 1 tera = 10^{12}). Fosse anche cento volte inferiore, questa quantità sarebbe pur sempre molto superiore a quella di una RAM (*read only memory*), che è la memoria ad accesso veloce di un moderno computer da tavolo (0.1-0.5 gigabyte) [ma sarebbe tuttavia confrontabile con quella, ad accesso assai più lento, di un disco fisso (0.1-0.5 terabyte)]. La differenza sostanziale tra il cervello e il computer sta dunque nell'enorme differenza di rapporto tra la quantità di memoria e la frequenza di lavoro, $\sim 10^{10}$ nel cervello umano, ~10 nel calcolatore. Questo significa che il nostro cervello può leggere e scrivere ad ogni ciclo di macchina un insieme di dati un miliardo di volte superiore di quello di un sistema seriale. Questo indica anche in modo molto efficace dove sta la differenza tra i processi paralleli e quelli seriali.

Come si spiega l'enorme differenza nell'utilizzo delle risorse tra la macchina parallela e quella seriale? Negli anni recenti l'elettronica e l'informatica hanno cercato di affrontare in modo sistematico, tanto nella teoria quanto nella pratica, il problema della costruzione di calcolatori paralleli. Le difficoltà che si sono incontrate testimoniano quanto l'informatica sia ancora lontana dal farci comprendere quali processi avvengano nei sistemi nervosi centrali degli animali e degli umani in particolare.

Una macchina seriale carica e scarica dati mediante alcune linee d'ingresso e d'uscita che trasmettono i dati come sequenze di bit. Essa è in grado di effettuare calcoli di complessità enormemente superiore a quella della sua struttura. Il comportamento ricorsivo esibito da una macchina di Turing universale, in corrispondenza a particolari dati iniziali, può risultare tanto complicato da non essere rappresentabile sinotticamente come un sistema di relazioni spaziali entro un insieme di dati. La macchina parallela lavora invece con insiemi molto numerosi di dati che sono trasmessi simultaneamente attraverso i suoi apparati d'ingresso e d'uscita. Di questo tipo sono i dati acquisiti dai sistemi sensoriali e quelli inviati all'apparato muscolare.

Può una macchina seriale effettuare un'elaborazione equivalente a quella di una macchina parallela contando sulla sua maggiore velocità? Può una macchina parallela effettuare processi ricorsivi di complessità equivalente a quella di una macchina seriale contando sulla sua maggiore complessità? Quale dei due tipi di macchine è in grado di fornire le prestazioni migliori per l'implementazione della capacità introspettiva?

Con opportune procedure di scansione un insieme finito di possibilità sincroniche può essere facilmente ordinato temporalmente e convertito in sequenze temporali di bit. In questo modo una macchina seriale è in grado di accettare in ingresso e produrre in uscita flussi di dati paralleli.

Il problema da risolvere, per rendere la macchina seriale capace di competere con quella parallela, sta nell'eseguire in tempi ragionevoli una sequenza numerosa d'operazioni su sottoinsiemi parzialmente sovrapposti di dati simultaneamente presentati in ingresso. Di questo tipo è ad esempio il problema che si presenta nella compressione dell'informazione visiva contenuta in un film. Per raggiungere alti coefficienti di compressione i sottoinsiemi di dati da elaborare devono essere relativamente numerosi e la percentuale di sovrapposizione deve essere elevata. Questo procedimento moltiplica i dati da elaborare per un fattore proporzionale alla numerosità dei sottoinsiemi rendendo spesso proibitivi i tempi di calcolo. Per diminuire i tempi di calcolo si può sfruttare il fatto che l'informazione che si ottiene elaborando un sottoinsieme è in gran parte già contenuta nell'informazione che si ottiene elaborando gli altri sottoinsiemi. Per riutilizzare quest'informazione di riporto, a mano a mano che è prodotta, l'algoritmo deve essere reso più complesso e avvalersi di procedure ricorsive complicate che richiedono tempi dei calcoli addizionali. Comunque la macchina, poiché opera sequenzialmente, non è in grado di utilizzare l'informazione di riporto non ancora prodotta. Insomma, la macchina seriale deve eseguire un grandissimo numero d'operazioni ricorsive complicate per ottenere quello che una macchina parallela potrebbe forse fare in un solo ciclo di funzionamento per semplice interazione simultanea delle sue parti.

D'altro canto, il problema fondamentale da risolvere per rendere la macchina parallela capace di competere con quella seriale nello svolgimento di un processo ricorsivo molto complicato è quello di suddividere il processo in sottoprocessi semplici e indipendenti; di distribuire in modo ottimale questi sottoprocessi tra le unità di calcolo; di integrarli e sincronizzali in un processo complessivo. In generale è assai difficile, se non addirittura impossibile, organizzare spazialmente un insieme molto grande di relazioni ricorsive in modo da rendere sincroniche le operazioni che compongono un processo seriale. La traduzione di una serie d'operazioni in un unico sistema di operazioni simultanee è possibile in casi particolari, ad esempio, quando la mancata esecuzione di un'operazione della serie non abbia conseguenze sulla corretta esecuzione delle altre, cioè nel caso che le operazioni siano commutative. Se le operazioni non sono commutative, il modo migliore di eseguirle è di eseguirle in successione temporale.

L'informatica dei processi paralleli identifica, in pratica, il calcolatore parallelo con un certo numero di calcolatori seriali che si scambiano messaggi (Codenotti e Leoncini, 1990). Il nostro interesse è rivolto invece ai sistemi d'elaborazione dell'informazione in cui la parallelizzazione è massima compatibilmente con la condizione della massima efficienza. E' plausibile che i principi teorici che devono esser posti alla base del calcolo parallelo siano diversi da quelli su cui si è basata la teoria del calcolo seriale. Cercheremo di impostare il problema senza pretendere di arrivare ad una soluzione. Nel seguito intenderemo per calcolatore parallelo un calcolatore i cui processi di calcolo sono massimamente parallelizzati.

Una caratteristica dei calcolatori seriali è la netta separazione dell'hardware dal software; della macchina che esegue le operazioni dai dati contenuti nei diversi sistemi di memoria. La struttura di un processo di calcolo dipende sia dall'architettura della CPU, che è fissa, sia dai programmi, che fanno parte dei dati immagazzinati nella memoria e sono variabili. La differenza tra i programmi e i dati da processare sta nei modi con cui questi diversi tipi di dati sono usati dalla CPU. La distinzione tra hardware e software si pone alla base della teoria del calcolo seriale perché essa corrisponde alla richiesta d'ottimizzazione dei processi seriali.

Per contrasto, nei calcolatori paralleli massimamente parallelizzati il software è necessariamente diffuso nell'hardware. Un grandissimo numero d'unità di calcolo concorre simultaneamente allo svolgimento di uno stesso processo complessivo, la cui efficienza dipende fortemente dall'architettura del sistema. Calcoli diversi richiedono architetture diverse e l'architettura ottimale per un certo tipo di calcolo non lo è affatto, in generale, per un altro. La corrispondenza tra i programmi di calcolo e le strutture ottimali di connessione interna è tale che è possibile affermare che i programmi dei calcolatori paralleli consistono essenzialmente della loro architettura interna.

Si sarebbe così portati ad immaginare che la soluzione migliore per un sistema di calcolo parallelo sia la costruzione di un complesso estesamente integrato d'unità di calcolo riccamente dotato di connessioni interne prestabilite. Le aree periferiche del sistema nervoso centrale, che sembrano principalmente deputate al filtraggio e alla riorganizzazione dei flussi dell'informazione sensoriale, sembrano soddisfare a questo requisito. Se questo fosse vero per tutte le aree e i nuclei del cervello, dovremmo concludere che il cervello non è altro che un gran filtro d'informazione.

Questo significherebbe che il cervello non può effettuare processi ricorsivi, perché questi richiedono un'incessante riorganizzazione del programma e quindi dell'architettura. Chiaramente, la plasticità delle connessioni neuronali ha un ruolo essenziale nel permettere la generazione di tali processi. D'altronde il cervello deve necessariamente eseguire processi ricorsivi, non solo perché questi sono indispensabili per la generazione delle capacità introspettive, ma anche, e in primo luogo, perché l'organizzazione dell'apprendimento e quella del comportamento richiedono in modo essenziale la ricorsività. Lo stesso "circuito" sensomotorio attraverso il quale avviene l'interazione col mondo esterno obbliga il cervello alla ricorsività. Di fatto, il cervello possiede dei formidabili circuiti paralleli deputati al processamento ricorsivo dell'informazione: il circuito di Papez e quello extrapiramidale sono i due più evidenti (Nieuwenhuys *et al.*, 1978).

Se poi richiediamo che il calcolatore parallelo sia universale, la capacità di riconfigurare rapidamente e incessantemente la sua architettura diventa indispensabile. Un calcolatore parallelo universale non può assumere un'architettura ottimale solo per un particolare tipo di processi. Esso deve essere abbastanza versatile da eseguire una gran varietà di processi disposti in sequenze imprevedibili; dunque deve essere in grado di modificare in breve tempo l'architettura di tutte le sue parti in modo reciprocamente coordinato. Questo significa in particolare che il cambiamento del modo di funzionamento di una parte deve influire sul cambiamento del modo di funzionamento delle parti connesse a monte e a valle rispetto alla direzione del processo effettuato. In altri termini, la condizione della ricorsività richiede che i flussi d'informazione parallela tra le diverse parti del sistema siano reciprocati da flussi d'informazione parallela che procedono in senso inverso. E' un dato di fatto che le connessioni tra tutte le parti operative del cervello sono reciprocate da connessioni inverse (Asanuma e Crick, 1986).

Parallelizzare al massimo un processo di calcolo - ciò che in generale potrà farsi in diversi modi - significa organizzarlo in modo che siano eseguiti simultaneamente quanti più calcoli sono possibili riducendo al minimo indispensabile i procedimenti che necessitano di ordinamento temporale. Questo richiede che il processo sia decomposto in una sequenza di fasi, in ognuna delle quali siano processati in parallelo quanti più sottoprocessi serialmente indipendenti e/o commutativamente interdipendenti siano possibili. Alla necessaria decomposizione temporale del processo in una successione di fasi, dovrà corrispondere una decomposizione spaziale in una successione di stadi, ciascuno costituito di molte unità di calcolo affiancate. La parallelizzazione appare tanto più conveniente per i cervelli animali se si considera, ad esempio, che l'acquisizione di dati dal mondo esterno, le letture e le scritture da e su memorie interne, il filtraggio e la discriminazione delle componenti diversamente significative di un flusso d'informazione, si presentano immediatamente come flussi d'informazione organizzati in parallelo. Si può comprendere come le funzioni e l'organizzazione ottimale delle unità di calcolo di un medesimo stadio possano dipendere in modo essenziale dalle proprietà generali dei flussi d'informazione entranti, in particolare dal grado d'omogeneità delle sezioni dei flussi, dalle loro forme di ridondanza e dall'organizzazione della loro complessità, nonché dalle funzioni alle quali sono destinati i flussi uscenti, dalla molteplicità delle destinazioni e dagli ordinamenti temporali richiesti.

La sincronizzazione delle modificazioni di stato delle unità locali è un'altra condizione indispensabile per il funzionamento del calcolatore parallelo. Un processo massimamente parallelizzato è caratterizzato dal fatto che lo stato del sistema a un certo istante è definito come uno stato globale in cui tutte le unità di calcolo hanno raggiunto una configurazione d'equilibrio temporaneo che contiene una distribuzione ben ordinata d'informazione. Se così non fosse gran parte del flusso d'informazione andrebbe perduta a causa del disordine spaziotemporale. Ne segue che l'intero processo di calcolo deve essere ordinato da forme di attività di sincronizzazione globali. Del resto questo corrisponde perfettamente a ciò che accade nel cervello.

Il trattamento dei flussi paralleli richiede molto spesso che le informazioni trasferite dalle varie parti del sistema entrino in relazione tra loro in qualche stadio successivo del processo. A tale fine, tutte le unità locali di un medesimo stadio iniziale devono trovarsi nella condizione di far convergere i risultati del loro funzionamento su una o più unità di uno stadio successivo. Se tra le varie componenti significative di un flusso d'informazione parallela devono stabilirsi relazioni complesse, allora, in generale, l'informazione veicolata dalle unità di calcolo di uno stadio iniziale deve convergere sincronicamente su una o più unità di qualche stadio più avanzato. Viceversa, se l'architettura degli stadi iniziali deve essere ottimizzata in modi finalizzati al buon funzionamento dell'intero processo, e quindi a quello degli stadi successivi, bisogna che sugli stadi iniziali converga informazione sincronizzata proveniente dagli stadi successivi.

Riassumendo, sulla base di semplici considerazioni di ottimalità funzionale abbiamo stabilito che un calcolatore massimamente parallelizzato deve possiede: una memoria distribuita e dispersa nelle unità di processamento; un'architettura variabile temporalmente diretta da attività oscillatorie generalizzate dell'intero sistema; sottosistemi deputati all'elaborazione sincronizzata dei flussi d'informazione; connessioni reciprocate tra i sottosistemi deputati allo svolgimento di processi ricorsivi. Tutte queste proprietà si ritrovano perfettamente implementate nei cervelli degli animali.

Ora, considerando che la gödelizzazione di un processo computazionale richiede potenze e velocità di calcolo che crescono esponenzialmente con la quantità d'informazione elaborata, e dunque fisicamente impossibili da raggiungere in tempi utili da parte di un calcolatore seriale, come si spiega che essa è invece possibile per un calcolatore massicciamente parallelo, sebbene relativamente lento come il cervello umano? Come si spiega che il cervello dispieghi di fatto una potenza di calcolo incomparabilmente superiore a quella di un calcolatore seriale di equivalente memoria e potenza operazionale?

Padova, 25.4.04

Bibliografia

1. Aiello, M., Albano, A., Attardi, G. e Montanari, U. (1976) *Teoria della computabilità, logica, teoria dei linguaggi formali*. Materiali Didattici ETS, ETS Pisa.
2. Arbib, M. (1968) *La mente, le macchine e la matematica*., Ed. Boringhieri, Bologna.
3. Asanuma, A. e Crick, F. (1986) in *Parallel Distributed Processing*, McClelland e Rumelhart Eds., Vol. 2, 333-371.
4. Barlow, H.B. (1968) in *Cybernetics*, Evans e Robertson Eds., p. 183-207, Butterworths, London.
5. Bartlett, S.J. (1992*) Reflexivity. A Source-Book in Self-Reference*. North-Holland.
6. Boole, G. (1976) *Indagine sulle leggi del pensiero*. Einaudi Ed. Torino.
7. Codenotti, B. e Leoncini, M. (1990) *Fondamenti di calcolo parallelo*. Addison-Wesley.
8. De Luca, A. e Ricciardi, L.M. (1981) *Introduzione alla cibernetica*. F.Angeli Ed., Milano. Mazurkievicz, A. (1977), voce *Algoritmo* in *Enciclopedia*, Einaudi Ed., Torino.
9. Hebb, D.O. (1949) *The Organization of Behavior*. John Wiley Ed.; traduzione italiana (1975) *L'organizzazione del comportamento*, Ed. F.Angeli.
10. Hofstadter, D.R. (1985) *Gödel, Escher, Bach: un'Eterna Ghirlanda Brillante*. Adelphi Ed. Milano.
11. Konolige. K. (1985) *A computational Theory of Belief Introspection*. Proceedings IJCAL-85. Anche in *Reflexivity*, Bartlett Ed. (1992), pp: 343-362.
12. Libet, B., Wright, E.W.Jr., Feinstein, B. and Pearl, D.K. (1979) Subjective referral of the timing for a conscious sensory experience. *Brain*. 102:193-224.
13. McCulloch, S.W. & Pitts, W. (1943) *Bull. Math. Biophysics* , 5:115-133.
14. Von Neumann, J. (1966) *Theory of Self-Reproducing Automata*. University of Illinois Press, Urbana.
15. Nieuwenhuys, R., Voogd, J. e van Huijzen, Chr. (1978) *Sistema nervoso centrale. Testo-Atlante.* Piccin Editore, Padova.
16. Nobili, R. e Pesavento, U. (1996) Generalized von Neumann Automata. In *Artificial Worlds and Urban Studies*, Besussi e Cecchini Ed.s. Convegni DAEST, Dipartimento di Analisi Economica e Sociale del Territorio, Venezia.
17. Putnam, H. (1979) voce *Logica* in *Enciclopedia*, Einaudi Ed., Torino.
18. Shannon, C. (1971) *La teoria matematica delle comunicazioni*. Ed. Eta Compas.
19. Tarski, A. (1966) Verità e dimostrazione. In *Verità e dimostrazione*. *Questioni di matematica*. A cura di C.Ciliberto, Lettura da Le Scienze, (1978) Le Scienze S.p.A editore, Milano.