

Giuseppe Trautteur

UNDICI TESI
SULLA SCIENZA COGNITIVA



Adelphiana

www.adelphiana.it

29 marzo 2002

Il termine «cognitivo» è superato, come lo sono i termini «cibernetica» e «intelligenza artificiale». Lo studio separato di percezione, deduzione, apprendimento, e così via – cioè della cognizione –, non si consolida in una unità disciplinare ben definita e dalla crescita bene indirizzata. Risulta quindi opportuno aprire l'indagine – come di fatto sta accadendo – ad altri, più problematici aspetti della mente: emozione, processi inconsci, intenzionalità, attitudini, dolore. In una parola: coscienza.

Negli ultimi cinquant'anni vari assalti sono stati sfermati alla conquista della mente, assalti che in molti casi hanno assicurato ricchi bottini collaterali, ma senza raggiungere la meta principale. Il tratto comune di questi tentativi è l'interdisciplinarietà introdotta dalla cibernetica. Questa interdisciplinarietà si configura come una sorta di alleanza tra discipline e saperi dalla improbabile convivenza, che a tutt'oggi rimane allo stato incoativo. Ne fanno parte la filosofia – e in particolare la filosofia della men-

te e la filosofia del linguaggio –, la linguistica, le neuroscienze – coacervo la cui recente legittimazione può considerarsi un esito davvero felice –, la psicologia, la logica, l'informatica, l'intelligenza artificiale, la scienza cognitiva, appunto – che a leggere le definizioni potrebbe ritenersi l'erede della cibernetica, esauritasi per la mancanza di risultati propriamente «cibernetici» e dissipatasi in molti rivoli di carattere o biologico (la prima rivista del campo, «Kybernetik», da anni ha preso atto di questa tendenza cambiando il suo nome in «Biological Cybernetics») o sociale, spesso compenetrandosi con la sistemica –, e infine la fisica, sia in proprio, sia come grande madre nel cui seno tutto nasce e tutto ritorna.

Vari tentativi all'interno dell'accademia europea e americana di fondere in un'unica, credibile disciplina – o almeno in una federazione coesa – questa alleanza hanno avuto vita breve, evolvendo rapidamente in istituzioni nell'ambito dell'informatica, sola vera disciplina nuova inopinatamente emersa da questa temperie. Tuttavia, pur avendo procurato strumenti utilissimi all'assalto sferrato cinquant'anni fa, l'informatica presenta problemi molto rilevanti e molto singolari. Innanzitutto, non siamo certi della sua identità culturale. È una scienza della natura? E se sì, qual è il suo oggetto, l'algoritmo? Oppure dobbiamo considerarla una disciplina dell'ingegneria? In secondo luogo, i suoi strabilianti effetti sociali rendono complessa una valutazione serena dei suoi metodi e dei suoi risultati.

Va detto che nei documenti del CNR le scienze co-

gnitive, tutte neurologiche o psicologiche – di cui la scienza cognitiva *simpliciter* è certamente la capofila –, vengono definite nei seguenti termini: «Le Scienze Cognitive rappresentano un approccio interdisciplinare allo studio del comportamento e dei processi mentali individuali e sociali...». Per parte sua, il MIUR caratterizza gli obiettivi formativi qualificanti della Classe di Lauree in Scienze Cognitive come: «... approfondite conoscenze di natura teorica e operativa per l’approccio interdisciplinare allo studio della mente e del comportamento degli organismi...».

Nei fatti, poi, ho l’impressione che oggi in Italia la scienza cognitiva sia molto vicina alla – o si sviluppi negli ambienti della – psicologia, principalmente cognitiva, *de facto* di ascendenza freudiana, e sia solo secondariamente legata all’intelligenza artificiale funzionalista (o *nouvelle* o *grounded* o neurale che sia), o ad aspetti teorici, fisici, biologici, filosofici dell’elaborazione simbolica. In ambiti locali, o sghembi rispetto all’obiettivo principale – la mente che comprende –, sono stati raggiunti risultati di notevole interesse. Peraltro, forse a causa del suo incoercibile fascino intellettuale, l’area tematica che sottende le discipline coinvolte nell’assalto alla mente è diventata di moda anche senza aver raggiunto risultati veramente propri e innovativi, quanto piuttosto ammantandosi dei grandi e reali successi dell’informatica e della neurobiologia. In particolare, nell’ultimo decennio è diventata socialmente accettabile nel mondo scientifico – ed è successivamente esplosa anche all’esterno – la pubbli-

cistica sulla coscienza. Libri sull'argomento – dalla raccolta di inappuntabili saggi scientifici a liete divagazioni d'opinione – escono dai torchi al ritmo di decine e decine all'anno. Nessuno ha veri risultati da comunicare, ma ci si affanna a ripetere che il problema della coscienza (spesso in consorzio con l'indagine fisica estrema) è il più importante e il più profondo di tutti, e che per la sua soluzione occorreranno idee interamente nuove, se pure. Dopodiché passano a esporre, come niente fosse, la loro teoria della coscienza.

Ma di una cosa siamo sicuri: nel complesso, nonostante le molte opinioni divergenti, si può dire che non si sa cosa sia la coscienza, o anche solo cosa possa significare il saperlo. Pure, ogni tanto, qualche passo avanti si registra. Tale va considerata ad esempio l'osservazione di James secondo cui la coscienza non è una «roba» (*stuff*), ma un processo. A pensarci bene, si tratta di un'osservazione sovvertitrice, perché significa realmente qualcosa solo in un contesto culturale in cui si sia disposti a considerare un moto, un cambiamento come qualcosa di sostanziale, e quindi in certo modo permanente, benché non necessariamente in senso temporale. (Anche da questa sponda, oltre che da quella della fisica, sembra imporsi una revisione della metafisica. Si noti come tale prospettiva, apparentemente contraria al comune buon senso, diventi più accettabile in ambito relativistico, nel quale processi come quello coscienziale di James possono essere assimilati alle linee d'universo nello spaziotempo. Tra l'altro, parliamo di idee nate anch'esse attorno al 1905; si vedano sotto, i commenti al par. 2).

Quello che si può fare, allora, è indagare le condizioni alle quali un'entità che si dica cosciente dovrebbe sottostare. Consideriamo l'introspezione e la riflessione. Normalmente si ritiene che la riflessione – cioè quella sorta di presenza a se stessi che consiste nell'agire, pensare o percepire sapendo che si sta pensando, agendo, percependo (ma si veda anche poco oltre) – sia un gradino verso la coscienza vera e propria. Tentandone un'indagine puramente neurologica, Damasio definisce questo stadio «coscienza nucleare», mentre definisce «autobiografica» la coscienza dispiegata nella quale si compie l'introspezione consapevole. Ma l'introspezione si compie a condizioni *meno* stringenti di quelle che rendono possibile la riflessione. Se un pensiero è un processo nel quale possiamo identificare un contenuto, ad esempio un puntatore a un altro processo o elemento di conoscenza, un pensiero *introspettivo* è un processo il cui contenuto è un processo che si sta svolgendo nel sistema medesimo, o che vi si è svolto (il contenuto è allora una traccia mnestica, il che rende contigue l'introspezione e il richiamo dalla memoria).

Certamente il processo introspettivo non è di semplice realizzazione, tuttavia esiste già, ed è persino commercializzato, in quegli artefatti computazionali – i sistemi operativi – che danno conto del proprio stato o esplicitamente si modificano in vista di uno scopo in un ambiente che varia.

Ben diverso il caso della riflessione, che la nostra esperienza fenomenologica ci suggerisce essere un

monitoraggio disatteso, anche se richiamabile a piacere, del processo primario (se esperisco consapevolmente il cielo azzurro non sto intrattenendo il pensiero di star facendo esperienza del cielo azzurro: quella sarebbe introspezione, magari cosciente, ma diversa dall'esperienza primaria del cielo azzurro). Pertanto il processo riflessivo deve essere costruito in modo che nella sua attività occorra integralmente il processo stesso. Diversamente dalla soluzione gödeliana, nella quale l'ente autoreferenziale contiene non se stesso, ma un nome di se stesso (ciò che risulta essere possibile attraverso la funzione di sostituzione e la numerazione gödeliana), nel caso della riflessione: a) il nome solo non basta, perché l'esperienza del nome e l'esperienza del processo medesimo sono due cose ben diverse, e presumibilmente l'esperienza si realizza, o consiste, nel fluire del processo; b) l'eventuale ricorso alla valutazione del nome – rivivere il processo – porta a una ricorsione non terminante. Benché certe patologie coattive mettano in evidenza lo scatenarsi di una ricorsione incontrollata, normalmente ciò non accade e quindi si rimane ingabbiati da una parte dalla necessità di un auto-inglobamento totale – una *mise en abîme* completa e infinita che neanche Escher è mai riuscito a realizzare –, dall'altra dalla conseguente violazione del principio di identità, in quanto un ente materiale non può essere identico a una sua parte.

Si badi che tutto ciò non vuole presentarsi come una teoria della coscienza: spero di non peccare di quel peccato che ho denunciato qualche riga sopra.

Anche se in ambito algoritmico questo problema potesse venir trattato (si vedano le osservazioni alla fine dei commenti al par. 8), rimarrebbe pur sempre il passo più lungo: come e perché l'esperienza si realizza, o consiste, in processi che nella prospettiva scientifica e filosofica attuale appaiono materialmente impossibili e metafisicamente contraddittori? (Propongo di leggere in questa chiave i passi della sesta *Meditazione* che portano Cartesio a concludere, mancandogli la conoscenza degli algoritmi come entità autonome, che la mente non può essere *extensa*).

2

Il nodo di tutti gli assalti al mentale sferrati da una base fisico-algoritmica (nonché di quelli neurobiologici) risiede nel problema della soggettività consapevole. Gli altri trattamenti del mentale presuppongono una connivente familiarità con esso.

Sembra che la mente sia immateriale. Questa concezione pare ripresentarsi in ogni cultura, a cominciare da quella occidentale. Si considerino la *res cogitans* del dualismo cartesiano, l'anima della Scolastica, gli enti intenzionali come trattati dalle leggi psicofisiche recenti (Chalmers, Koenderink), e poi, in ordine sparso, la psiche, gli ordinamenti giuridici, la responsabilità, la libertà, le dichiarazioni dei diritti di questa o quella classe di entità – per esempio i primati non umani –, solo apparentemente e



saurite nella loro materialità. Nella sfera dell'immateriale vengono fatti rientrare, d'abitudine, enti inerti o passivi, cioè pensati: concetti, universali, forme astratte, numeri e così via, insieme a enti attivi, cioè pensanti: anime, spiriti, coscienza, angeli, Dio. (C'è qui una qualche assonanza tra gli «immateriali inerti» e i contenuti del ben noto *Mondo3* di Popper. Se poi gli «immateriali attivi» assomigliano agli abitanti del *Mondo2* di Popper è assai

più discutibile). Gli immateriali inerti non sembrano esistere in forma autonoma, e le loro asserite capacità causali possono sempre essere sciolte con attribuzioni di causalità al parlante/pensante che li ha menzionati o ad altri enti materiali connessi, sintatticamente o referenzialmente, al nome del-

l'immateriale stesso. Per fare solo un esempio, «la sua bellezza lo affascinò» viene sciolto nelle reazioni neurobiologiche scatenate dall'analisi, presumibilmente visiva, della forma della donna. La parola «bellezza» è un mero codice linguistico per indicare una catena causale complessa, ma epistemicamente innocua. Naturalmente si tratta di un'operazione di codifica del tutto inconsapevole, il cui funzionamento risale a fasi animistiche, ma persistenti nella cultura della società – nella Grecia di Omero, secondo Onians, parole come *pneûma*, *phrén*, *thymós* denotavano sia parti o funzioni del corpo umano, sia le corrispondenti potenze immateriali.

Sappiamo che Snow parlava di due culture – ma la seconda, quella scientifica, in realtà non esiste. Esiste solo la cultura umanistica, la Kultur. Un letterato, un filosofo, un critico d'arte non può non essere colto, mentre per l'uomo di scienza essere colto è irrilevante (come pure lo è, stranamente, per l'artista visivo, o per il musicista). Perché? Perché nella scienza la fruizione in prima persona dei contenuti è superflua. Se un risultato scientifico viene percepito come «bello», oppure viene contemplato, la cosa sotto il profilo operativo e sociale è irrilevante. Non così in ambito umanistico, dove i cosiddetti valori, cioè le fruizioni in prima persona dei contenuti oggettivi, sono cruciali per l'economia sociale del milieu colto. E allora il punto non è fondere le due culture, ma raggiungere un obiettivo forse più modesto – recuperare alla Kultur i contenuti scientifici. Alcuni esempi. Il *Mann ohne*

Eigenschaften di Musil una caratteristica, almeno una, ce l'ha: è un matematico. C'è chi fa critica d'arte usando non solo le forme nello spazio delle immagini, ma anche quelle nello spazio coniugato delle frequenze. Zeki da anni lavora a un'interpretazione dell'esperienza estetica visiva in termini neurocognitivi. Senza una formazione scientifica non si possono gustare appieno le non poche pagine in cui Gadda dimostra di essere, innanzitutto, un ingegnere.

La cultura umanistica è strutturata in modo tale da non dover tematizzare la differenza tra prima e terza persona: la distinzione certo non le è estranea, ma non ne tocca i gangli vitali. I due punti di vista convivono senza attriti distruttivi. Non così per la scienza, che nasce e per quasi tre secoli si sviluppa esclusivamente in terza persona. I primi, grandiosi scostamenti da questo regime avvengono contemporaneamente nel 1905, quando in meccanica quantistica e nella relatività la prima persona entra nella descrizione in terza persona, mentre la psicoanalisi promuove l'indagine in terza persona della fenomenologia della prima persona. Pochi anni dopo emerge l'evidenza che la crisi dei fondamenti della matematica dipende in modo cruciale da una situazione analoga, e il lavoro di Gödel permette, con le ben note conseguenze, l'introduzione di una sorta di punto di vista in prima persona nello studio dei sistemi logico-algoritmici (codificando nel metalinguaggio il sistema dall'interno del sistema stesso). Naturalmente, non è questa la sede per dar conto delle complesse relazioni tra autorefe-

renzialità, riflessione e rapporti in prima e terza persona (si vedano anche i commenti al par. 1). Ora, l'esistenza autonoma di enti immateriali attivi è il nodo tuttora irrisolto. La connivenza cui si è fatto cenno risiede nel presupporla non, ovviamente, come fatto esperienziale insopprimibile, ma come una datità conoscitivamente accertata. Se ci si pensa, Cartesio e Husserl hanno voluto affrontare proprio questa esigenza di rigore intellettuale.

3

Tutto quello che accade nel mentale dipende dall'esistenza prioritaria di una qualche soggettività. La situazione è però molto diversa da quella che si presentò nelle fasi iniziali della meccanica quantistica.

È chiaro che mentre la mente può studiare molti enti diversi, fra tutti gli enti solo la mente può studiare la mente. L'asimmetria non è banale quanto appare al senso comune, perché in linea teorica potrebbero esistere varie specie del genere mente: umana, di primate non umano, angelica, di mammifero, artificiale, consapevole, risvegliata (nel senso dell'ascesi indiana), risvegliata (nel senso dell'*awareness* anglosassone), e così via, e alcune di queste specie potrebbero conoscerne alcune altre, ma non tutte, come nei casi di personalità multiple. Inoltre, alcune specie potrebbero non esistere o coalescere nella mente umana – l'unica, francamente, di cui fino ad ora si abbia notizia. Qui non intendiamo

rivolgere l'attenzione allo studio delle «altre menti», né chiederci se la mente artificiale sarà funzionalmente omologa a quella umana o no. Il punto che ci interessa è che quando l'oggetto di studio di una mente è un'altra mente o addirittura la medesima sorgono circolarità preoccupanti (cui abbiamo già accennato nel par. 1).

Nella definizione di informazione proposta, seriamente, da von Weizsäcker («Informazione è ciò che viene capito oppure ciò che viene prodotto dall'informazione») ritroviamo sia il circolo vizioso della ricorsione non terminante, sia il riferimento alla soggettività: sta qui l'origine dei problemi dell'autoriferimento e della tematizzazione del metalinguaggio che percorrono tutto il Novecento.

La millenaria questione se l'idea sia *ante rem* o *post rem* certo non sarà decisa qui. Nonostante quanto asserito nel par. 2 sullo scioglimento degli immateriali inerti, alcuni *a priori* sembrano ineluttabili. Per esempio la contemplazione dell'aritmetica lascia perplessi sulla possibilità della naturalizzazione dei numeri naturali.

La via qui auspicata è quella di una naturalizzazione metodologicamente inesorabile e al tempo stesso memore del buon senso. Per esempio i teorici dell'identità, o altre varietà di eliminativisti, forse non hanno lavorato abbastanza a delineare in modo intelligibile cosa potrebbe essere un mondo mentale totalmente privo di soggettività.

Ci si trova come di fronte a un koan: considerare il soggetto come la datità primaria donatrice di sen-

so e anche come la meta di un cammino di spiegazione che si fonda sulla materia.

Fenomeni analoghi furono dibattuti nella prima meccanica quantistica e ripresi oggi, in parte anche per la possibilità di eseguire esperimenti difficoltosi. Ma l'interpretazione coscienziale del cosiddetto collasso del pacchetto d'onde nel processo di misurazione, sostenuta per esempio da Pascual Jordan o E. Wigner e dibattuta nella vasta letteratura sul gatto di Schrödinger, oggi appare marginalizzata. Certamente una grande acquisizione della fisica fu la scoperta fattuale della necessità, che *a priori* è di natura metodologica e non fattuale, di considerare il processo di misurazione alla stregua del processo in esame. Ma la difficoltà sembra risiedere piuttosto nella transizione dalla coerenza di fase del mondo microscopico alla statistica termodinamica del mondo macroscopico. E nel perché solo il secondo sia accessibile all'esperienza diretta. Purtroppo confusione e superficialità hanno portato alcuni autori a sfruttare erroneamente l'analogia epistemica tra meccanica quantistica e coscienza, con esiti imbarazzanti (si veda anche il par. 9).

4

È errore fatale confondere lo studio del mentale presupponendo la soggettività con lo studio del mentale in quanto soggettività. Si tratta di uno spartiacque intrinsecamente discontinuo. Studiare il mentale ignorando la soggettività equivale a studiarlo presupponendola, ancorché inconsapevolmente.

Si ripropone qui il tema della «connivenza». La distinzione la cui mancanza qui si lamenta è, di fatto, la mossa iniziale, e ormai familiare, di ogni procedimento analitico: sono coinvolto anch'io nel dominio dell'analisi? (ne abbiamo un esempio in corso nella cosiddetta politica italiana). Per quanto detto sopra la risposta, in questo contesto, è certamente sì, e quindi vanno prese opportune precauzioni. Per esempio, appena si abbandona il procedurale per abbracciare una «tecnica» denotazionale, si introduce surrettiziamente una soggettività nell'artefatto che si sta studiando: come se la semantica esistesse al di fuori di una comunità di soggetti che capiscono consapevolmente! Tutti gli sforzi di naturalizzazione dell'informazione, della conoscenza, dell'intenzionalità a me noti falliscono o, banalmente, per difetti di argomentazione, oppure per petizione di principio, perché – con un'onestà e una chiarezza spesso autodistruttive – vi si presuppone, con Cartesio e Brentano, un *aliud quid* oltre la materia.

5

I considerevoli risultati raggiunti in anni recenti dal «cognitivo puro» sono stati ottenuti in una situazione di prudente ed esplicita «epoché» della soggettività consapevole: questo è il presupposto che rende impossibile lo sviluppo del trattamento della consapevolezza in tale direzione.

Il termine *epoché* che qui compare richiama il lavoro filosofico, profondamente affine all'assalto al

mentale menzionato più sopra, che portò Husserl alla riduzione fenomenologica. Per motivi che non sono in grado di sondare, è accaduto che la deriva antiscientifica del movimento fenomenologico e il suo ancorarsi alla nozione di intenzionalità abbiano condotto in primo luogo alla dicotomia tra filosofia analitica e filosofia continentale e, per quel che qui ci concerne, all'inimicizia tra la fenomenologia (Dreyfus, Searle e altri) e qualsiasi attacco al mentale su base fisico-algoritmica. Eppure alcune analisi fenomenologiche recenti, ma anche quelle di Kohler, della *Gestalt*, di Merleau-Ponty, sembrano tendere verso le stesse mete dell'intelligenza artificiale o della neurobiologia. Basta riflettere, per esempio, alla formulazione del problema della percezione visiva nel «*binding problem*», che, com'è noto, tematizza la difficoltà di capire come gli elementi sensoriali, spezzettati in numerosissimi rivoli analitici dal sistema nervoso centrale (SNC), vengano poi ricomposti – dove? come? da chi? – in una percezione unitaria. C'è un'incomprensione di fondo originata, forse, dall'ignoranza dei reciproci risultati e dalla sottolineatura prematura, eccessiva e ostentata dei due rispettivi caposaldi: l'oggettività del simbolo fisicamente realizzato (calcolabilità effettiva, Physical Symbol System Hypothesis di Newell e Simon) e il primato dell'intenzionalità che, sotto il travestimento delle leggi psicofisiche, sfocia in una cosmologia animistica. Questi due caposaldi sono la manifestazione colta del fenomeno prima persona/terza persona, che appare perfettamente gestibile nel contesto del senso comune

e della vita quotidiana, ma presenta quelle interdipendenze antinomiche in cui identifichiamo uno degli aspetti sinora impenetrabili della soggettività.

6

Nella maggior parte degli scienziati-cognitivi l'infatuazione per il calcolatore nasconde l'ignoranza persino di quel poco che si conosce dell'elaborazione simbolica e dell'algoritmo in informatica.

Siamo qui in presenza di uno di quegli effetti sociali non dominati, scaturiti dall'informatica, cui si alludeva nei commenti al par. 1. Il calcolatore si impone come il fine dei mezzi, se così si può dire. In altri termini, per qualunque scopo sia necessaria una elaborazione simbolica – o spesso anche in assenza di necessità – si ritiene che tale elaborazione debba essere compiuta con il calcolatore (il cosiddetto «computer»). Tuttavia, il calcolatore stesso – in linea con i prodotti della tecnologia avanzata – non si fa dominare da coloro che credono di usarlo, e ne sono invece largamente usati. Il problema, affrontato in letteratura meno estesamente di quanto sarebbe auspicabile, è quello dell'*homo faber* di fronte al suo artefatto.

7

Questa ignoranza, come pure la scarsa familiarità con la scienza di base – la fisica –, svuota di fatto la cosid-

detta interdisciplinarietà. E pone un problema di strategia didattica per lo sviluppo della ricerca futura sulla coscienza.

Le neuroscienze che – giustamente – propongono di coordinare diverse discipline biologiche nello studio integrato del processo mentale non si sono aperte ad altri campi disciplinari. In generale manifestano diffidenza verso il trattamento omogeneo del biologico e dell'inanimato, e addirittura ostilità verso la nozione di algoritmo, in diretto contrasto con la loro ispirazione cibernetica.

Esiste una posizione di prima persona «sociale» che consiste nel partecipare alle votazioni «credendoci», nel prendere per vero il contenuto esplicito del contratto sociale o, in presenza di un *minimum* di scetticismo, nel considerarlo auspicabile, doveroso e così via. Ovviamente le cose non stanno così, ma c'è a dir poco una forte riluttanza a guardare realmente il mondo sociale ed economico, per tacere di quello politico, in terza persona. Pochi l'hanno fatto – certamente Marx e, di recente, Bateson, con il suo uso singolare dell'ideologia cibernetica. Questa temperie stinge su tutte le scienze cosiddette «soft», e il suo influsso è particolarmente nocivo su medicina e biologia.

Pregiudizio e diffidenza non fanno predisporre l'opportuna didattica, la cui mancanza genera a sua volta ignoranza e diffidenza. Il problema è analogo a quelli, per esempio, dello sviluppo autonomo del Terzo Mondo o del superamento degli integralismi cristiani e islamici. Sorge inopinata-

mente, in ambito scientifico, il problema della diversità. La diversità è la base del razzismo o la meta della tolleranza? (Aporia della tolleranza: non può/deve tollerare l'intolleranza). E che ruolo può avere la diversità in un luogo in cui (ci) si (illude di) persegu(ir)e la verità? Come mai nell'intelligenza artificiale, negli studi sulla coscienza, nella scienza cognitiva, sussistono scuole di pensiero diverse – eliminativista, connessionista, misteriana, simbolico-computazionalista, A-Lifista – tutte volte a indagare e denunciare come e perché le altre sbaglino e perché, attenzione, il loro metodo sia quello giusto? È auspicabile la creazione di lauree specialistiche e dottorati che, a partire da una base scientifica generica (matematica, logica, informatica, fisica, biologia), costruiscano l'interdisciplinarietà all'interno dei singoli individui, anche in mancanza di una disciplina unificante. A questo proposito, non posso dimenticare la mia formazione prima nel Gruppo di Cibernetica, creato da Eduardo Caianiello presso l'Istituto di Fisica teorica dell'Università di Napoli – dove chi piantava microelettrodi nel tetto ottico delle rane comunicava con chi scriveva quadruple di macchine di Turing o regolava con potenziometri le connessioni di reti neuroniche –, e poi presso il programma in Communication Sciences della University of Michigan, presto «evoluto» in Dept of Computer Science, dove mi furono fatte studiare linguistica, psicologia sperimentale, metamatematica, teoria dell'informazione. Esempi di quei tentativi di breve durata menzionati all'inizio.

L'elaborazione simbolica è insieme strumento e argomento di studio. Si studia la mente insieme alla natura dell'algoritmo. Le divisioni oggi in voga fra simbolico, neurale, sistemico, e così via, derivano da una lettura superficiale che sceglie di ignorare il problema ancora insondato del fondamento materiale dell'algoritmo.

Nelle formule dei manuali (il Rogers, tanto per fare un nome) l'algoritmo si presenta come un'entità matematica ben definita, con istruzioni, ordinamenti, esecuzioni... Ah, esecuzioni. In un algoritmo i collegamenti del mentale con il fisico sono assai più complessi di quelli, pure assai oscuri, del numerale con il numero o del simbolo con la cosa simboleggiata, perché l'algoritmo non è solo materiale – il problema della circoscrizione dell'effettività tematizzato nella Tesi di Church-Turing –, ma attivo. Un momento: è attivo l'algoritmo, che in fondo dovrebbe essere un ente astratto, o il programma, indubitabilmente materiale? In apertura del suo libro *Vision*, e introducendo la sua interpretazione del processo visivo, Marr propose di distinguere fra l'algoritmo in senso astratto o intuitivo e le sue possibili realizzazioni materiali nei programmi – distinzione che sembra richiamare quella tra enunciato e proposizione. Ma nell'ambito della semantica dei linguaggi di programmazione, la semantica denotazionale indica come ente di riferimento dell'oggetto «scrittura programma» una funzione, e non l'algoritmo di cui la «scrittura programma» doveva essere presumibilmente un «nu-

merale». Proposta operativa: considerare le locuzioni «algoritmi», «programma» e «procedura effettiva» come sinonimi che denotano un'entità dalla fenomenologia straripante, molto vicina al mondo mentale ma radicata nella materia. Rientrerebbero così negli algoritmi i programmi Lisp, le CPU, le api bottinatrici quando eseguono o interpretano la loro tipica danza, le reti neurali di *Aplysia*, il software in genere.

Il radicamento nella materia va poi considerato sotto due profili. Da una parte si devono demarcare gli enti materiali che sono (sede di) procedure effettive (per irridere il funzionalismo, ameni filosofi hanno «dimostrato» che una parete muraria è un automa finito); dall'altra occorre definire le proprietà fisiche di tali enti in termini di dissipazione energetica, reversibilità, e così via, che è ciò di cui si occupano Landauer, Bennett e altri. E qui va decisa la natura della materia algoritmica: è continua o discreta? Nel caso delle reti neuroniche artificiali, che Kleene ha dimostrato essere algoritmi in una certa loro varietà originaria, si è ad esempio pensato che la continuità ne arricchisse la natura.

E se l'ipotesi algoritmica della mente deve sopravvivere, un problema in merito alla coscienza, non ancora tematizzato e analogo a quello dell'unità percettiva o «*binding problem*», è posto dalla percezione del primo dei continui fenomenologici (qui nel senso proprio della filosofia fenomenologica), cioè del tempo. Forse, nello spirito di Bergson, la coscienza si può identificare con la percezione del presente: ma che cosa significano esattamente frasi di questo genere?

Mentre fin dai tempi della Scolastica la materia è principio di individuazione, il problema dell'identità di un algoritmo è tutt'altro che chiaro. Le copie di un programma pongono lo stesso tipo di problema, e di risposte al problema, dei libri e delle opere d'arte (soprattutto dopo le indagini di Benjamin sulla loro riproducibilità), ma gli accessi ricorsivi (è sempre l'attività dell'algoritmo ad arricchire le cose), o le invocazioni di un oggetto, o la semplice rilocalizzazione virtuale di una procedura pongono interrogativi al tempo stesso banali e inquietanti. Banali perché alla fine tutto viene risolto con l'uso di pile. Inquietanti perché vi compare seriamente la nozione di livello, in genere fuorviante in quanto usata come una sorta di panacea per molti enigmi metafisici (si pensi, per esempio, alla recente invenzione filosofica della supervenienza). Ora, definire cosa sia letteralmente un livello, al di qua del suo valore metaforico – tranquillizzante perché fondato sulla nostra dimestichezza con il campo di gravità –, è generalmente difficile quanto la situazione «spiegata» dall'uso della parola livello. La problematicità del concetto di identità per la procedura effettiva potrebbe essere un buon avvio alla comprensione della coscienza – questa strana cosa in cui ciò che pensa è identico e distinto da ciò che viene pensato.

rati tentativi di procacciarsi un'illusione di senso senza affrontare il nodo della coscienza. Il nuovo fascino per la materia occulta una fame di spirito, presente anche altrove nella società, che però si piega a soddisfacimenti banalmente consumistici, come la New Age. La soggettività è legata al sistema nervoso centrale e solo al sistema nervoso centrale. Il cui collegamento con il mondo – dalla corteccia somatosensitiva verso l'esterno – è fungibile. La struttura del sistema nervoso centrale, e soprattutto la forma dell'elaborazione del segnale-simbolo che vi si svolge, non lo sono.

In ambito neuropsicologico circola da alcuni anni la metafora – ma si tratta piuttosto di una falsa analogia – che propone di considerare il cervello come l'hardware di cui la mente sarebbe il software. L'analogia [*Es*] *ist nicht einmal falsch*, non è nemmeno erronea, come sembra dicesse il grande Pauli delle teorie fisiche dilettantesche e cervellotiche. Infatti la distinzione tra hardware e software, tra macchina fisica e programma, è lungi dall'essere chiarita, e i suoi confini sono mobili: basta riflettere a questa distinzione nel caso di un'architettura banale con CPU e RAM rispetto ad architetture più esotiche e riconfigurabili, per esempio le FPGA. Una possibile interpretazione di tale apparente dicotomia è quella degli invarianti adiabatici nel senso di Ehrenfest, o dei moti secolari in meccanica celeste, oppure ancora della dinamica seconda, cioè dei parametri invece che delle variabili libere nella teoria delle catastrofi. Alcuni sistemi fisici possono essere descritti in (almeno) due riprese suc-

cessive, e ciò rende assai più intelligibile il sistema rispetto a tentativi direttamente globali. L'esempio più noto è quello di due corpi gravitanti in presenza di un'altra interazione. In questo caso è facile determinare il moto dei due corpi nel sistema di riferimento del loro centro di massa e in assenza dell'altra interazione. Successivamente li si considera concentrati in questo centro di massa, e si studia il loro comportamento rispetto all'altra interazione. (Sole, Terra, Luna. Dapprima Terra/Luna come sistema isolato, poi Sole/(Terra+Luna)). L'orbita finale si ottiene come l'orbita della Luna attorno alla Terra che si sposta lentamente attorno al Sole (lo studio diretto del sistema a tre corpi, com'è noto, non si dà). Tutto ciò è comunque possibile solo se la velocità dei due corpi sulle loro orbite è molto maggiore di quella del loro centro di massa, cosicché il moto di quest'ultimo «perturba» di poco la forma delle orbite.

Ora, l'hardware sarebbe la parte della dinamica a variazione più lenta, il software quella rapida, di superficie. L'esecuzione di un programma sarebbe assimilata all'orbita del pianeta a parametri dinamici congelati, quella dell'hardware – per esempio l'evoluzione del repertorio di istruzioni di macchina (transizione a macchine RISC, ecc.) – darebbe luogo a un fascio di esecuzioni assimilabili alle variazioni secolari delle orbite planetarie sotto diversi influssi perturbativi. Nella prassi informatica, è il problema della «*upward compatibility*».

Ma nel modesto arsenale matematico della teoria della calcolabilità non abbiamo strumenti potenti

e generali quanto quelli della fisica matematica, e anche questa interpretazione – che deriva da una «*Adiabatic learning hypothesis*» formulata da Caianiello nel 1961 per le reti neuroniche – va presa con molta circospezione. Il suo punto nodale è comunque la sottolineatura del fatto che, nella prima riduzione metafisica – quella contro cui *ex post* oggi molti si scagliano pur tranquillamente vivendone –, hardware e software sono entrambe entità materiali. E dal momento che non si sa con precisione cosa sia un algoritmo, le loro differenti fenomenologie sono ardue da dipanare quanto quelle del cervello e della mente.

Eppure si continua a insegnare – credo a ragione – che il mondo della elaborazione simbolica è stratificato in livelli. In tale contesto, non a caso, si può dare una definizione accurata di livello in termini di macchina di Turing universale. Il più basso è l'hardware, quelli intermedi sono il sistema operativo, l'interprete, e così via (si vedano i commenti ai parr. 8 e 11), mentre il più alto è il top level che tutti conoscono. Le relazioni tra questi livelli sono omologhe tra loro e danno luogo al concetto di virtualità. Ciò potrebbe rivelarsi una chiave utile per l'interpretazione di quegli enormi spazi di lavoro (reminiscenze del lavoro di Baars?) che sono le cortecce associative. Non potrebbe essere interessante considerare il tessuto nervoso sede di processi che sarebbero gli «hardware virtuali» appropriati alla – straordinariamente diversificata – attività cerebrale?

Evitando analogie semplicistiche, i neuropsicologi

di cui sopra eviterebbero anche il riferimento, assai più ambiguo e deleterio, alla sussistenza di una dualità mente-spirito, che immancabilmente compare quando si propongono spiegazioni del genere *obscurum per obscurius*. (Un altro luogo tipico di queste dubbie procedure si riscontra nelle elucubrazioni sulla mente alla Penrose, o alla Hameroff, che chiamano in causa la meccanica quantistica o l'inesistente teoria della gravità quantistica).

I tentativi di costruire senso ancorando ambigualmente le strutture algoritmico-simulative – per loro natura effettive, quindi materiali (si vedano i parr. 8 e 9) – una seconda volta nella materia sono un regresso considerevole nel processo di chiarificazione della soggettività. Una dimestichezza men che superficiale con la fisica teorica svela infatti la profonda oscurità del concetto di materia.

Il senso si ha solo nella «*middle distance*» ontologica, per usare una frase suggestiva di Brian C. Smith. Harnad, Brooks e altri ancora incappano in una *petitio principii* usando nello smontaggio analitico del senso comune un materiale concettuale tipico del senso comune, la cui saldezza addirittura proviene dalla solidità esperienziale della materia, già largamente contestato fin dall'antichità classica e il cui superamento costituisce uno dei successi del moderno. Un'altra possibile donazione di senso è quella inerente alla fenomenologia husserliana, che per essersi allontanata dallo spirito scientifico costituisce tuttavia un tassello affascinante, ma non immediatamente integrabile nel mosaico delle conoscenze sulla mente.

La materia invocata nel *grounding* sembra essere un principio fondante, che riaffiora da qualche cosmologia «alla Giordano Bruno» per contrapporsi all'«algoritmo». (C'è qui qualche analogia con l'elevazione della materia biologica di cui fa uso Searle). Ciò che a noi interessa non è tanto la materia come principio fondante, ma quella comune materia di cui è interessante conoscere come si comporta nel SNC, e non nell'ambiente – ambiente al quale in ogni caso il SNC risulta somatotopicamente connesso, e che è da esso (SNC) «rappresentato». Be', non rappresentato, perché a parte microelettrodi, PET, SQUID o fMRI, non c'è nessuno che guardi l'attività materiale del SNC e ne possa fruire come se «rappresentasse» l'ambiente. Il quale ambiente, peraltro, è largamente coinvolto nel funzionamento del SNC attraverso le codifiche neurosensoriali e neuromotorie.

10

Con poche eccezioni, che riguardano lo studio della percezione e dell'elaborazione visiva, la neurobiologia continua a ignorare come i segnali neurali vengano elaborati. Tutt'al più prende in considerazione l'intensità della «segnalazione» neurale e, spesso appiattendolo lo studio dinamico su quello anatomico, non immagina neppure la smisurata ricchezza che può separare una procedura dalla sua esecuzione. Ciò è causato dall'ignoranza o dalla diffidenza istituzionali – in quanto determinate dalla struttura dell'attuale processo formativo – della e verso l'informatica e, più in generale, l'elaborazione simbolica.

27

Dalla nascita della cibernetica, verso la fine degli anni Quaranta, a oggi, nessun centro di ricerche ha mai intrapreso un progetto di dimensioni veramente vaste. In compenso, spesso si è lamentata l'inettitudine della tecnologia rispetto alla raffinatezza della «natura», pur nella constatazione di una crescita imprevista delle capacità tecniche, che oggi permetterebbero la costruzione di macchine confrontabili quantitativamente con il sistema nervoso centrale. Ma perché non sono stati intrapresi progetti simili a quelli per acceleratori di particelle, la mappatura del genoma, le sonde spaziali, il telescopio Hubble, o i quarantennali – e peraltro fallimentari – tentativi di ottenere la fusione termonucleare controllata? Per mero terrore, come ha finalmente detto Minsky in un messaggio a una lista in rete l'11 ottobre scorso? Questo progetto avrebbe dovuto tentare la realizzazione, con qualunque metodologia, di una entità artificiale percipiente e senziente. Sarebbe stato l'unico modo per rendere maturo il cosiddetto metodo simulativo, a tutt'oggi alquanto privo di riscontri sperimentali.

Che risorse ingenti siano state investite nell'informatica e nelle sue applicazioni non legate all'industria, al commercio o all'intrattenimento è cosa ovvia. Per fare solo un esempio di applicazioni «interne» e scientifiche, si pensi agli strabilianti risultati nel trattamento dell'immagine visiva, alle svariate forme di tomografia, oppure alle capacità dei motori di ricerca, che sembrano talvolta superare le capacità di associazione mentale dell'uomo. E,

in chiave un po' diversa, alla nascita di un atteggiamento intenzionale – alla Dennett, ma reale – verso le macchine, assieme al manifestarsi nei programmi di tratti, voluti o «accidentali», nettamente psicologici: tutti sottoprodotti del lavoro in Human Computer Interaction.

Il rammarico è piuttosto rivolto alla mancanza di una ricerca concentrata sulla conferma di ipotesi, teorie o intuizioni ancora fluide, come invece accade, con più o meno successo, negli esempi citati sopra. Se ne sente parlare da sempre, ma appunto soltanto parlare. La sensazione di chi ha alle spalle quattro decenni di frequentazione e meditazione nel campo è quella di una specie di Gioco dell'Oca nel quale, in modo inopinato ma sistematico, si viene ogni volta retrocessi alla casella n. 1, e si è costretti a ricominciare tutto daccapo.

L'attuale ripresa dalla casella n. 1, dove si passa dalla locuzione «intelligenza artificiale» a quella «scienza cognitiva», trae speranza da parole come *neurocomputing*, *softcomputing*, agenti intelligenti, *situatedness*, A-Life, XML, *grounding*, dall'abbandono, spesso non motivato, di altre parole quali funzionalismo, Soar, retroazione, Information processing psychology, e dalla promozione – quantomeno sul piano tecnologico – di altre ancora, come *theorem proving*, rappresentazione della conoscenza, percezione e, infine, emozione. I meccanismi che collegano – o secondo Damasio identificano – quest'ultima con la coscienza nucleare sfuggono però all'indagine e alla simulazione. Francamente non si sa da dove cominciare a costruire un arte-

fatto come quello la cui mancata realizzazione ho appena lamentato. Forse la vera causa di questa stasi è l'assoluta mancanza anche solo di un punto di partenza per affrontare l'abisso che divide la coscienza, sede della prima persona, dal resto del mondo. Per tacere della percezione del dolore, a proposito della quale l'indagine algoritmico-neurologico-funzionalista è ancora ferma a Dennett. La soggettività o coscienza vive poi una sua vita ricchissima di parole, come accennato all'inizio, in un suo mondo filosofico e neurologico senza contatti, anzi quasi ostile all'informatica.

In realtà, se volessi riassumere questi quattro decenni di assalti alla mente, direi che il processo conoscitivo si sta svolgendo autonomamente, senza un progetto euristico consapevole. La «crescita impreveduta delle capacità tecniche» appena menzionata non si è certo realizzata da sola. È l'esito, oltre che naturalmente dell'interesse mercantile, anche della seguente mossa, più volte ripetuta nel mondo informatico: quando non si sa fare qualcosa si passa a preparare un ambiente generico (qui, come anche nella nozione di virtualità, si avvertono disattese ombre teoriche delle numerazioni gödeliane), nel quale, come non è chiaro, sarà più facile farlo. Da mosse simili sono nati negli anni Cinquanta i linguaggi ad alto livello e poi i sistemi operativi, i sistemi di gestione delle basi di dati, la programmazione logica e così via. È da notare che si tratta di mosse diverse dal semplice – si fa per dire – miglioramento di un sistema dato. Ad esempio nel settore «hard» – sì, ci sono differenze tecnologiche

tra «hard» e «soft» – il cablaggio di una intera macchina universale (CPU) su un solo chip o il raddoppio della memoria per area di wafer ogni tanti anni sono «semplici» miglioramenti. Viceversa, all'unico sgambetto mai tentato nel settore «hard», le strutture parallele, è toccata una storia di successo analoga a quella del reattore a fusione termoneucleare controllata.

Molti risultati di questa crescita volta alla sostituzione artificiale delle attività mentali o alla loro semplificazione e al loro incremento, realizzati a scopo di lucro, presentano un enorme interesse conoscitivo. In altri termini, il programma mai scritto dell'intelligenza artificiale si sta realizzando non grazie a sforzi intenzionali, ma come effetto collaterale del processo tecnico-industriale. Un fenomeno analogo si realizzò tra gli ultimi decenni del Settecento e i primi dell'Ottocento, quando lo sviluppo impetuoso delle macchine termiche precorse di molti anni la teoria termodinamica. Forse la storia si sta ripetendo?

Che la scienza cognitiva si «riappropri» dunque di questi risultati, servendosene per rilanciare, più consapevolmente, la sua indagine.

Attratto per natura dagli artefatti – orologi, ferrovie, congegni ottici e acustici – Giuseppe Trautteur ha compiuto studi di fisica. Attualmente, ma da sempre, lavora all'ipotesi algoritmica per discernere, se c'è, l'articolazione fra mente e materia.