

I testi proposti sono tratti da **A. Damasio "L'errore di Cartesio"** e si compongono di quattro passi, i primi due sono tratti dall'introduzione, il terzo presenta il caso di Phineas Gage, l'ultimo discute e critica il dualismo cartesiano.

Seguono poi due brani tratti da H. Simon "Le scienze dell'artificiale" e la voce "**Intelligenza artificiale**" (di **R. Cordeschi**) dell'Enciclopedia filosofica.

INTRODUZIONE

Non saprei dire con certezza che cosa abbia acceso il mio interesse per le basi neurali della ragione, anche se so bene quando arrivai a convincermi che le tradizionali opinioni sulla natura della razionalità potevano non essere corrette. Sin da giovane ero stato avvertito che le decisioni solide scaturiscono da una mente fredda, e che emozioni e ragione non si mescolano di più che olio e acqua. Così, ero cresciuto nella consuetudine di pensare che i meccanismi della ragione fossero disposti in una provincia separata della nostra mente, nella quale non doveva consentirsi alle emozioni di penetrare. E quando cominciai a pensare al cervello che sta dietro a quella mente, immaginai sistemi neurali distinti per la ragione e per l'emozione – secondo una credenza ampiamente diffusa sulla relazione tra le due, in prospettiva sia mentale sia neurologica.

Ma ora mi si parava davanti agli occhi l'essere umano intelligente più freddo e meno emotivo che si potesse immaginare, la cui ragione pratica, però, era talmente menomata da produrre, nelle vicende della vita quotidiana, un seguito di errori, una per-

petua violazione di quel che voi e io riterremmo socialmente appropriato e vantaggioso dal punto di vista personale. La sua mente era stata del tutto sana fino a che un danno neurologico non colpì un particolare settore del suo cervello, provocando da un giorno all'altro una profonda deficienza della capacità di decidere. Gli strumenti che si è soliti giudicare necessari e sufficienti per il comportamento razionale erano intatti: disponeva della memoria, dell'attenzione e della conoscenza richieste; il suo linguaggio non mostrava peccche; egli poteva eseguire calcoli, e poteva affrontare la logica di un problema astratto. Alla mancanza di capacità decisionale si accompagnava solo una vistosa alterazione della capacità di provare sentimenti. Sentimenti alterati e una ragione imperfetta si presentavano assieme come conseguenze di una specifica lesione cerebrale, e questa correlazione mi suggeriva che il sentimento fosse una parte integrante del modo di operare della ragione. Due decenni di lavoro clinico e sperimentale, condotto su un gran numero di pazienti neurologici, mi hanno permesso di ripetere questa osservazione più e più volte, e di trasformare un indizio in un'ipotesi verificabile.¹

Ho cominciato a scrivere questo libro volendo proporre l'idea che la ragione può non essere così pura come la maggior parte di noi ritiene che sia, o vorrebbe che fosse; che i sentimenti e le emozioni possono non essere affatto degli intrusi entro le mura della ragione: potrebbero essere intrecciati nelle sue reti, per il meglio e per il peggio. Sia nell'evoluzione sia in ogni singolo individuo, le strategie della ragione umana probabilmente non si sono sviluppate senza la forza guida dei meccanismi di regolazione biologica dei quali emozione e sentimento sono espressioni notevoli. Per di più, anche dopo che le strategie del ragionamento si sono assestate, negli anni della formazione, il loro effettivo dispiegamento dipende in larga misura dalla ininterrotta capacità di provare sentimenti.

Con questo non si nega che l'intervento delle emozioni e dei sentimenti possa causare grande disordine nei processi di ragionamento. La saggezza tradizionale ci ha insegnato che ciò può avvenire, e anche analisi recenti del processo di ragionamento normale rivelano l'influenza potenzialmente dannosa delle inclinazioni emotive. È, quindi, ancor più sorprendente e nuovo che l'assenza di emozione e sentimento sia non meno dannosa, non meno capace di compromettere la razionalità che ci rende peculiarmente umani e ci permette di decidere in armonia con un senso di futuro personale, di convenzione sociale e di moralità.

Né ciò equivale a dire che sono i sentimenti, quando esercitano un'azione, a decidere per noi, o che noi non siamo esseri razionali. Io suggerisco soltanto che certi aspetti del processo dell'emozione e del sentimento sono indispensabili per la razionalità. Nei casi migliori, i sentimenti ci volgono nella direzione giusta, ci conducono al luogo appropriato di uno spazio decisionale nel quale possiamo fare bene operare gli strumenti della logica. Quando dobbiamo formulare un giudizio morale, decidere sull'andamento di una relazione personale, scegliere come evitare di trovarci privi di mezzi nella vecchiaia, o semplicemente decidere che cosa fare degli anni che verranno, siamo di fronte all'incertezza. Emozione e sentimento, insieme con i processi fisiologici che ne costituiscono la base nascosta, ci assistono nello scoraggiante compito di prevedere un futuro incerto e di pianificare in sintonia le nostre azioni.

Prenderò le mosse da un'analisi del celebre caso di Phineas Gage, vissuto nel secolo scorso, che per la prima volta rivelò una connessione tra razionalità menomata e uno specifico danno cerebrale; poi esaminerò gli studi recenti sulle sue controparti moderne, e passerò in rassegna le relazioni tra i risultati della ricerca neuropsicologica condotta su esseri

umani e di quella condotta su animali; inoltre, proporrò la tesi che la ragione umana dipende da diversi sistemi cerebrali, operanti di concerto attraverso molti livelli di organizzazione neuronica, anziché da un unico centro. Nel farsi della ragione cooperano sia le regioni cerebrali di livello «alto» sia quelle di livello «basso», dalle cortecce prefrontali all'ipotalamo e al midollo allungato.

Nell'edificio neurale della ragione, i livelli più bassi sono gli stessi che regolano l'elaborazione delle emozioni e dei sentimenti, insieme con le funzioni somatiche necessarie per la sopravvivenza dell'organismo. A loro volta, questi livelli mantengono relazioni dirette e mutue con pressoché tutti gli organi del corpo; questo viene così posto direttamente all'interno della catena di operazioni che generano le conquiste più alte del ragionamento, della decisione e, per estensione, del comportamento sociale e della creatività. Emozione, sentimento, regolazione biologica hanno tutti un ruolo nella ragione umana. I livelli più modesti del nostro organismo fanno parte del ciclo della ragione superiore.

È interessante ritrovare l'ombra del nostro passato evolutivo al livello più squisitamente umano della funzione mentale, anche se già Charles Darwin prefigurò l'essenza di questo risultato quando scrisse dell'impronta indelebile delle origini inferiori che gli esseri umani portano nella loro struttura corporea.² Tuttavia la dipendenza della ragione alta dal cervello inferiore non trasforma la prima in ragione bassa. Il fatto che l'agire in accordo con un principio etico richieda la partecipazione di connessioni semplici nel nucleo encefalico non immiserisce il principio stesso. L'edificio dell'etica non crolla, la moralità non ne è minacciata e, nell'individuo normale, la volontà rimane volontà. Quello che può cambiare è il nostro modo di vedere come la biologia abbia contribuito all'origine di certi principi etici che scaturiscono in un dato contesto sociale,

-D Per quanto sulle prime possa sorprendere, la mente esiste dentro e per un organismo integrato: le nostre menti non sarebbero quello che sono se non fosse per l'azione reciproca di corpo e cervello - nel corso dell'evoluzione, durante lo sviluppo dell'individuo e nel momento presente. La mente dovette prima essere per il corpo, o non sarebbe potuta essere. Sulla base del riferimento che il corpo fornisce con continuità, la mente può allora avere a che fare con molte altre cose, reali e immaginarie.

Quest'idea si radica sui seguenti enunciati: 1) il cervello umano e il resto del corpo costituiscono un organismo non dissociabile, integrato grazie all'azione di circuiti regolatori neurali e biochimici integranti (che includono componenti endocrini, immunitari e nervosi autonomi); 2) l'organismo interagisce con l'ambiente come un insieme: l'interazione non è del solo corpo né del solo cervello; 3) i processi fisiologici che noi chiamiamo «mente» derivano dall'insieme strutturale e funzionale, piuttosto che dal solo cervello: soltanto nel contesto dell'interagire di un organismo con l'ambiente si possono comprendere appieno i fenomeni mentali. Il fatto che l'ambiente sia, in parte, un prodotto dell'attività stessa dell'organismo semplicemente sottolinea la complessità delle interazioni che bisogna tenere in conto.

Quando si parla di cervello e di mente, non è consuetudine fare riferimento agli organismi. Di fronte all'evidenza che la mente scaturisce dall'attività dei neuroni, si discute solo di questi, come se il loro funzionamento potesse essere indipendente da quello del resto dell'organismo. Ma via via che stu-

diavo i disturbi della memoria, del linguaggio e della ragione, presenti in numerosi esseri umani colpiti da lesioni al cervello, sempre più mi si imponeva l'idea che l'attività mentale – nei suoi aspetti più semplici come in quelli più alti – richiede sia il cervello sia il resto del corpo. Quest'ultimo, a mio avviso, fornisce al primo più che un puro sostegno e una modulazione: esso fornisce la materia di base per le rappresentazioni cerebrali.

È un'idea sostenuta da fatti; vi sono ragioni che la rendono plausibile e ragioni per le quali sarebbe bello se le cose stessero davvero così. Fra queste ultime, soprattutto la considerazione che la precedenza del corpo qui suggerita potrebbe gettar luce su una delle più tormentose domande che assillano gli esseri umani da quando hanno cominciato a indagare sulla mente: come avviene che siamo coscienti del mondo attorno a noi, che sappiamo ciò che sappiamo, che sappiamo di sapere?

Nella prospettiva dell'ipotesi accennata, amore e odio e angoscia, qualità come gentilezza e ferocia, la soluzione pianificata di un problema scientifico o la creazione di un nuovo artefatto si basano tutti su eventi neurali all'interno di un cervello, purché questo sia stato e sia in interazione con il corpo cui appartiene. L'anima respira attraverso il corpo, e la sofferenza, che muova dalla pelle o da un'immagine mentale, avviene nella carne.

Ho scritto questo libro come se fossi in conversazione con un immaginario amico: un amico curioso, intelligente e avveduto, poco edotto di neuroscienze ma molto della vita. Con lui ho stipulato un accordo: la conversazione doveva essere vantaggiosa per entrambi. Egli avrebbe dovuto imparare qualcosa sul cervello e su quelle arcane cose che diciamo mentali; io avrei dovuto guadagnare in ispirazione mentre mi sforzavo di spiegare il che e il come, secondo me, del corpo, del cervello e della mente.

SCIAGURA NEL VERMONT

Phineas P. Gage

Siamo nel New England, alla fine dell'estate del 1848. Phineas P. Gage, venticinquenne caposquadra di un'impresa di costruzioni, sta per precipitare dalle stelle alle stalle. Un secolo e mezzo dopo, la sua caduta sarà ancora significativa.

Gage lavora per la Rutland & Burlington Railroad: a lui è affidata una squadra numerosa di operai (una *gang*, in gergo), con il compito di gettare i binari per una nuova linea ferroviaria che attraverserà il Vermont. Nelle due settimane precedenti, sono avanzati un po' a rilento, in direzione della cittadina di Cavendish; adesso sono impegnati su una sponda del Black River. Il loro lavoro è tutt'altro che agevole, per i numerosi affioramenti di roccia dura. Invece di aggirare ogni scarpata, si è deciso di far saltare la roccia, dove necessario, per aprire la via a un tracciato più diritto e piano. Gage sovrintende a tutti questi lavori, ed è magnificamente all'altezza; atletico e ben proporzionato nel suo metro e settanta di statura, si muove con rapidità e precisione. Somiglia a un giovane James Cagney, a un elegante Yankee Doodle che balli il tip tap su rotaie e traversine, vigoroso e aggraziato al tempo stesso.

Agli occhi dei suoi capi, poi, Gage è ben più che un dipendente in gamba come ce ne sono tanti: lo definiscono l'uomo « più efficiente e capace » tra quanti hanno assunto.¹ E questo è un bene, perché il lavoro richiede tanto maestria fisica quanto sottile concentrazione, soprattutto quando bisogna preparare le detonazioni. Qui occorre procedere in modo ordinato, passo dopo passo. Dapprima, va scavato un foro nella roccia; poi il foro va riempito per metà con esplosivo in polvere, si deve inserire la miccia e la polvere deve essere coperta con sabbia. Questa deve essere « pressata », cioè compattata con una accurata sequenza di colpi inferti con una barra di ferro. Infine, bisogna accendere la miccia. Se tutto va bene, la polvere esploderà dentro la roccia; la sabbia è essenziale, poiché senza la sua copertura protettiva l'esplosione si sfogherebbe fuori della roccia. Anche la forma della barra e il modo in cui viene usata sono molto importanti: Gage se ne è fatta fare una apposta, dando indicazioni precise, ed è un virtuoso della pressatura.

Alle quattro e mezzo di questo caldissimo pomeriggio, Gage ha appena finito di introdurre la polvere esplosiva nel foro, e ha detto all'uomo che è con lui di coprirla con la sabbia, fino al riempimento. Qualcuno da dietro lo chiama e Gage si volta per un istante a guardare sopra la propria spalla destra. Si distrae e prima che l'operaio abbia versato la sabbia comincia a pestare con la barra di ferro, direttamente sulla polvere. Subito fa sprizzare scintille dalla roccia, e la carica gli esplode sul viso.²

L'esplosione è così violenta che la squadra si blocca, raggelata: ma in pochi secondi è chiaro quel che è successo. L'esplosione è stata diversa dal solito, e la roccia è rimasta intatta; diverso dal solito anche il rumore che l'ha accompagnata, un sibilo come di razzo scagliato nel cielo. Ma non sono fuochi d'artificio: è un assalto, una scarica. La barra metallica penetra nella guancia sinistra di Gage, fora la base

della scatola cranica, attraversa la parte frontale del cervello ed esce, velocissima, dalla sommità della testa, per andare a cadere, impiasticciata di sangue e di tessuto cerebrale, a una trentina di metri di distanza. Phineas Gage è stato scagliato a terra e giace stordito, nel chiarore del pomeriggio; muto, ma sveglio. E così sono tutti gli impotenti spettatori.

«Orribile incidente» intitolano – prevedibilmente – la notizia il «Daily Courier» e il «Daily Journal» del 20 settembre, una settimana più tardi. Il 22 settembre, il «Vermont Mercury» curiosamente intitola *Mirabile incidente*; il «Boston Medical and Surgical Journal» con più precisione sceglie il titolo *Passaggio di una barra di ferro attraverso la testa*. A giudicare dalla concretezza con la quale raccontano la vicenda, si direbbe che i giornalisti conoscessero bene i racconti straordinari e i racconti dell'orrore di Edgar Allan Poe. Forse è così, anche se non sembra verosimile: i racconti gotici di Poe non sono ancora popolari, e lo stesso Poe morirà un anno dopo, sconosciuto e squattrinato. O forse è solo che l'orribile è nell'aria.

Tutti si sorpresero che Gage non fosse rimasto ucciso all'istante; l'articolo medico della rivista di Boston riporta che «subito dopo l'esplosione il paziente fu rivoltato sulla schiena»; che poco dopo egli mostrò «alcuni movimenti convulsi delle estremità» e «nel giro di pochi minuti parlò»; che «i suoi uomini (dei quali egli era un beniamino) lo sollevarono e a braccia lo trasportarono fino alla strada, che distava solo poche pertiche (una pertica essendo pari a circa 5 metri) e lo posero a sedere su un carro trainato da buoi, sul quale egli – seduto con la schiena eretta – percorse più di un chilometro, fino all'albergo di Joseph Adams»; e che Gage «scese dal carro da solo, con un piccolo aiuto da parte dei suoi operai».

Vediamo un po' meglio la figura di Adams. Egli è il giudice di pace di Cavendish; inoltre possiede l'u-

nico albergo e l'unico spaccio di alcolici del paese. Io lo vedo più alto di Gage, largo quasi il doppio, apprensivo come la sua stazza da Falstaff lascia immaginare. Dopo essersi avvicinato a Gage, manda subito a chiamare il dottor John Harlow, uno dei medici di Cavendish. Nell'attesa, mi sembra quasi di sentirlo esclamare: « E allora, signor Gage, che cosa abbiamo? » e anche, perché no: « Ohimè, che disgrazia ci tocca vedere! ». Agita le mani, quasi ad allontanare ciò che è successo, e conduce Gage verso la zona ombreggiata del portico dell'albergo. Le descrizioni lo indicano come una *piazza*, che suona grande e spazioso e aperto; forse è grande e spazioso, ma non aperto: è giusto un porticato. Adams ora forse fa bere a Phineas Gage una limonata, o magari un bicchiere di sidro ben fresco.

Dall'esplosione è passata un'ora; il sole si sta abbassando e il caldo è più sopportabile. È in arrivo un altro medico, il dottor Edward Williams, un collega – più giovane – del dottor Harlow, il quale in seguito descriverà la scena con queste parole: « Quando lo vidi era seduto su una sedia nella *piazza* dell'albergo di Adams, a Cavendish. Appena mi avvicinai, mi disse: "Dottore, qui c'è lavoro per voi". Prima ancora di scendere dalla carrozza, avevo notato la ferita sulla sua testa: si potevano vedere chiaramente le pulsazioni del cervello. Notai anche qualcosa che non riuscì a spiegarmi, prima di esaminare la testa: la sommità di questa si presentava come un imbuto rovesciato. In seguito avrei scoperto che ciò era dovuto al fatto che l'osso attorno all'apertura era fratturato per una lunghezza di quasi 5 centimetri in tutte le direzioni. Ho dimenticato di precisare che l'apertura attraverso il cranio e i tegumenti aveva un diametro di quasi 4 centimetri: i bordi di questa apertura erano rovesciati, e nel complesso la ferita dava l'impressione che un oggetto sagomato a cuneo avesse attraversato la testa muovendo dal basso verso l'alto. Mentre io gli esa-

minavo la ferita, Gage raccontava ai presenti in che modo era stato colpito: parlava con tale lucidità ed era talmente desideroso di rispondere che io rivolsi le mie domande a lui piuttosto che agli uomini che erano presenti al momento dell'incidente e che ora ci attorniavano. Poi Gage mi riferì alcune delle circostanze, come ha poi fatto più volte, e io posso affermare con sicurezza che né allora né in una qualsiasi occasione successiva – salvo una – io lo considerai men che perfettamente razionale. L'unica volta in cui ne dubitai fu una quindicina di giorni dopo l'incidente, allorché insistette a chiamarmi John Kirwin – e però a tutte le mie domande rispose in modo corretto».³

Il fatto che Gage fosse sopravvissuto risulta tanto più sorprendente quando si considerino la forma e il peso della barra di ferro. Henry Bigelow, professore di chirurgia alla Harvard, la descrive con queste parole: « Il ferro che attraversò il cranio pesa 6 chilogrammi; è lungo 110 centimetri e ha un diametro di poco più di 3 centimetri. L'estremità che penetrò per prima è rastremata, per una lunghezza di 18 centimetri, e termina con una punta del diametro di circa 6 millimetri. A queste circostanze, forse, il soggetto deve la propria salvezza. Il ferro non somiglia ad alcun altro strumento, ed è stato fatto da un fabbro della zona seguendo le indicazioni del cliente ».⁴ Gage è piuttosto rigoroso per tutto quanto riguarda il suo lavoro e i ferri del mestiere.

Sopravvivere all'esplosione con una ferita al capo così ampia e profonda, essere capace di parlare e camminare e di mantenersi coerente subito dopo l'incidente: tutto ciò è ben sorprendente. Ma sarà altrettanto sorprendente che Gage superi il sopravvenire dell'inevitabile infezione. Il dottor Harlow conosce bene l'importanza della disinfezione. Non ha gli antibiotici, ma ricorrendo alle sostanze chimiche all'epoca disponibili egli pulisce la ferita energicamente e con metodo preciso; inoltre dispone il

paziente in posizione semisdraiata, in modo da facilitare il drenaggio. In Gage si produrranno forti febbri e almeno un ascesso, sul quale il bisturi di Harlow interverrà prontamente. Alla fine, la giovane età e la robusta costituzione di Gage avranno il meglio, con l'assistenza dell'intervento divino, come dirà Harlow: « Io l'ho medicato, Dio l'ha guarito ».

Phineas Gage sarà dichiarato guarito nel giro di meno di due mesi. Tuttavia quest'esito stupefacente impallidisce al confronto con la straordinaria svolta che la sua personalità sta per subire. Il suo carattere, i suoi gusti, i suoi sogni, le sue aspirazioni: tutti cambieranno. Il corpo di Gage può essere ben vivo e vegeto, ma c'è un nuovo spirito che lo anima.

Gage non era più lui

Quello che accadde esattamente possiamo arrivare a conoscerlo, oggi, a partire dal resoconto che il dottor Harlow stese vent'anni dopo l'incidente.⁵ È uno scritto degno di fede, che offre dovizia di fatti e un minimo di interpretazione; convincente dal punto di vista sia umano sia neurologico, ci permette di ricostruire non solo la figura di Gage, ma anche quella del suo dottore. John Harlow era stato un insegnante di scuola, prima di entrare al Jefferson Medical College di Philadelphia; quando prese in cura Gage aveva cominciato da pochi anni a esercitare come medico. Il caso di Gage divenne l'interesse dominante di tutta la sua vita, e sospettò che abbia fatto desiderare a Harlow di essere uno studioso – il che forse non era nei suoi programmi quando cominciò la pratica medica nel Vermont. Il successo riportato nel curare Gage e i resoconti preparati per i suoi colleghi di Boston furono i momenti di gloria della sua carriera, e dovette turbarlo la nube che incombeva sulla terapia di Gage.

Nel suo racconto, Harlow descrive come Gage

riacquistò le forze e come si ristabilì pienamente, dal punto di vista fisico: poteva toccare, udire, vedere, e non subì paralisi agli arti o alla lingua. Aveva perduto la vista dall'occhio sinistro, ma con il destro ci vedeva perfettamente. Camminava con passo fermo, usava le mani con destrezza e non mostrava impaccio nella parola o nel linguaggio. E tuttavia, come riferisce Harlow, «l'equilibrio, per così dire, tra la sua facoltà intellettuale e le sue disposizioni animali» era stato distrutto. Questi cambiamenti divennero manifesti non appena fu superata la fase acuta della lesione cerebrale. Ora egli era «bizzarro, insolente, capace a volte delle più grossolane imprecazioni, da cui in precedenza era stato del tutto alieno; poco riguardoso nei confronti dei compagni; insopportabile di vincoli o consigli che contrastassero i suoi desideri; a volte tenacemente ostinato, e però capriccioso e oscillante; sempre pronto a elaborare molti programmi di attività future che abbandonava non appena li aveva delineati ... Un bambino, nelle sue manifestazioni e capacità intellettuali, ma con le passioni animali di un adulto robusto». Il linguaggio è talmente osceno e degradato che alle donne si consiglia di non rimanere a lungo in sua presenza, o la loro sensibilità ne sarà turbata. Né gli ammonimenti più vigorosi dello stesso Harlow riescono a riportare il nostro sopravvissuto a un comportamento corretto.

Questi nuovi aspetti della personalità di Gage erano in acuto contrasto con le «abitudini moderate» e con la «grande forza di carattere» che gli erano state proprie, come si sapeva, prima dell'incidente. Gage aveva avuto «una mente assai equilibrata, e quanti lo conoscevano guardavano a lui come a un uomo abile e avveduto nei suoi affari, molto energico e tenace nel perseguire tutti i programmi d'azione che si fosse prefissi». Non c'è dubbio che egli avesse successo nell'ambito del suo lavoro e dei suoi tempi. Di fronte a un cambiamento così ra-

dicale, amici e conoscenti quasi non lo riconoscevano, e osservavano tristemente che « Gage non era più lui ». Quando si ripresentò, i suoi datori di lavoro non lo riassunsero, poiché « secondo il loro giudizio la sua mente adesso era cambiata in modo così marcato che non potevano ridargli il posto che aveva prima ». Il problema non stava in un difetto di abilità o di capacità fisica; il problema era il suo nuovo carattere.

La vicenda continua a seguire il proprio corso. Non più in grado di fare il caposquadra, Gage trovò lavoro in vari allevamenti di cavalli; ma si arguisce che in genere lo lasciava presto, per un suo capriccio, oppure ne veniva allontanato per indisciplina. Harlow osserva che egli era molto bravo a « trovare sempre qualcosa che non gli andasse bene ». Cominciò allora la sua carriera come attrazione da circo. Gage venne presentato nel Barnum's Museum di New York, in vanagloriosa esibizione delle proprie ferite e della barra di ferro. (Questa non lo abbandonava mai, dichiara Harlow, il quale fa notare il forte attaccamento di Gage a oggetti e animali. Era un comportamento nuovo e in qualche modo fuori dell'ordinario, che si potrebbe chiamare « comportamento del collezionista », e che io ho osservato in vari pazienti colpiti da lesioni come quella di Gage e anche in individui autistici).

Allora, e assai più di oggi, il circo traeva profitto dalla crudeltà della natura. La varietà endocrina comprendeva i nani, la donna più grassa del mondo, l'uomo più alto, quello con la mandibola più sviluppata; la varietà neurologica comprendeva i giovani con « pelle da elefante », cioè colpiti da neurofibromatosi – e ora Gage. Possiamo immaginarcelo, in tale pittoresca brigata, dedito a barattare la propria disgrazia con denaro.

Quattro anni dopo l'incidente, un altro *coup de théâtre*: Gage parte per l'America del Sud. Qui lavora forse in allevamenti di cavalli e talvolta come con-

duttore di diligenze a Santiago e a Valparaiso. Non si sa molto altro sulla sua vita da espatriato, se non che nel 1859 la sua salute andava peggiorando.

Nel 1860, Gage ritornò negli Stati Uniti per andare a vivere con la madre e la sorella, che nel frattempo si erano trasferite a San Francisco. Lavorò dapprima in una fattoria di Santa Clara, ma non vi rimase a lungo. Anzi, continuò a muoversi, trovando lavori occasionali come manovale nella zona. Risulta chiaro che non era una persona indipendente, e che non era capace di assicurarsi quel tipo di occupazione regolare e sufficientemente retribuita che aveva avuto una volta. La fine della caduta si approssimava.

La San Francisco del 1860 me la immagino come un luogo brulicante di traffici, frequentata da audaci imprenditori, in un fiorire di attività agricole, minerarie, armatoriali. Qui possiamo trovare la madre e la sorella di Gage, quest'ultima sposata con un agiato mercante della città (tal D.D. Shattuck, Esquire); a questo ambiente il Phineas Gage di una volta sarebbe potuto appartenere. Ma non è qui che lo troveremmo, se potessimo viaggiare all'indietro nel tempo. Lo troveremmo, invece, probabilmente, intento a bere, tra una rissa e l'altra, in un quartiere malfamato; e quando la faglia si muoverà, i minacciosi sussulti della terra non lo coglieranno, attonito come tutti, nel bel mezzo di una conversazione con i signori del commercio. Egli si era unito alla compagnia di quegli sconfitti i quali « erano venuti in California per morire », per dirla con le parole che avrebbe usato Nathanael West, alcuni decenni più tardi e qualche centinaio di chilometri più a sud.⁶

Gli scarsi documenti disponibili suggeriscono che in Gage si manifestassero attacchi di epilessia. La fine giunse il 21 maggio 1861, dopo un malessere che durò poco più di una giornata. Gage ebbe una violenta convulsione che gli provocò perdita di coscienza, e poi una serie di convulsioni meno gravi;

morì senza avere riacquistato conoscenza. Io penso che egli fu vittima di uno stato di male epilettico – condizione nella quale le crisi convulsive si fanno pressoché continue e portano alla morte. Aveva allora trentotto anni. Nessun necrologio comparve sui giornali di San Francisco.

Perché Phineas Gage?

Qual è il possibile significato di una vicenda così strana? La risposta è semplice: mentre altri casi di danno neurologico, avvenuti più o meno nella stessa epoca, rivelavano che il cervello era la base del linguaggio, della percezione e della funzione motoria, fornendo anche in generale elementi più pertinenti, la storia di Gage additava un fatto sbalorditivo: vi erano nel cervello umano – anche se non si sapeva precisarne il come – sistemi deputati al ragionamento più che a qualsiasi altra funzione, e in particolare alle dimensioni personali e sociali del ragionamento. Un danno cerebrale poteva comportare la fine dell'osservanza di regole etiche e convenzioni sociali acquisite in precedenza, anche quando né il linguaggio né l'intelletto sembravano compromessi. Senza volerlo, l'esempio di Gage indicava che qualcosa nel cervello aveva a che fare specificamente con proprietà peculiarmente umane, tra cui la capacità di anticipare il futuro e di pianificare in accordo con tale anticipazione, all'interno di un ambiente sociale complesso; il senso di responsabilità verso sé stessi e verso gli altri; la capacità di predisporre la propria sopravvivenza in modo deliberato, in ottemperanza al proprio libero volere.

L'aspetto che più colpisce, in questa triste vicenda, è il contrasto tra la struttura normale della personalità quale era prima dell'incidente e i nefandi caratteri che emersero dopo, e che sembra si siano mantenuti per il resto della vita di Gage. Egli aveva

saputo, una volta, tutto ciò che gli occorreva sapere sulle scelte da compiere per migliorarsi; aveva il senso di una responsabilità personale e sociale (che si rifletteva nei progressi della sua carriera), si curava della qualità del proprio lavoro, richiamava l'ammirazione di datori di lavoro e colleghi. Era ben adattato, in termini di convenzioni sociali, e la sua condotta appare marcata da un'etica. Dopo l'incidente, in lui non vi è più alcun rispetto per le convenzioni sociali; l'etica – nel senso più ampio del termine – veniva violata, le decisioni da lui prese non tenevano in conto il suo più vero interesse, ed egli si dava a inventare storie « che non avevano alcun fondamento se non nella sua fantasia », secondo le parole di Harlow. Non si vede alcun indizio di preoccupazione per il proprio futuro, né alcun segno di preveggenza.

Le alterazioni della personalità di Gage non erano modeste. Egli non poteva compiere buone scelte, e le scelte che faceva non erano neutre; non erano le timide, riservate decisioni di qualcuno la cui mente è stata menomata ed è timoroso di agire, ma erano decisamente svantaggiose. Si potrebbe provare a supporre che il suo sistema di valori ora fosse differente, oppure che, se era rimasto lo stesso, in nessun modo i valori di prima potessero influenzare le sue decisioni. Non vi è alcun elemento che possa dirci se è vera l'una o l'altra ipotesi, ma esaminando vari pazienti colpiti al cervello in modo simile a Phineas Gage mi sono convinto che nessuna delle due spiegazioni coglie ciò che davvero accade in tali circostanze. Parte del sistema di valori rimane, e può essere utilizzata, in termini astratti: ma è priva di legami con le circostanze della vita reale. Quando i Phineas Gage di questo mondo devono operare nella realtà, le conoscenze che avevano acquisite prima dell'incidente influenzano solo in minima misura il processo di decisione.

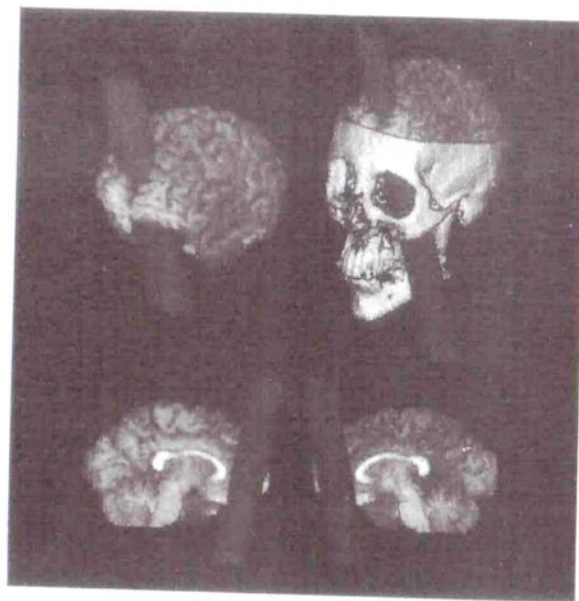
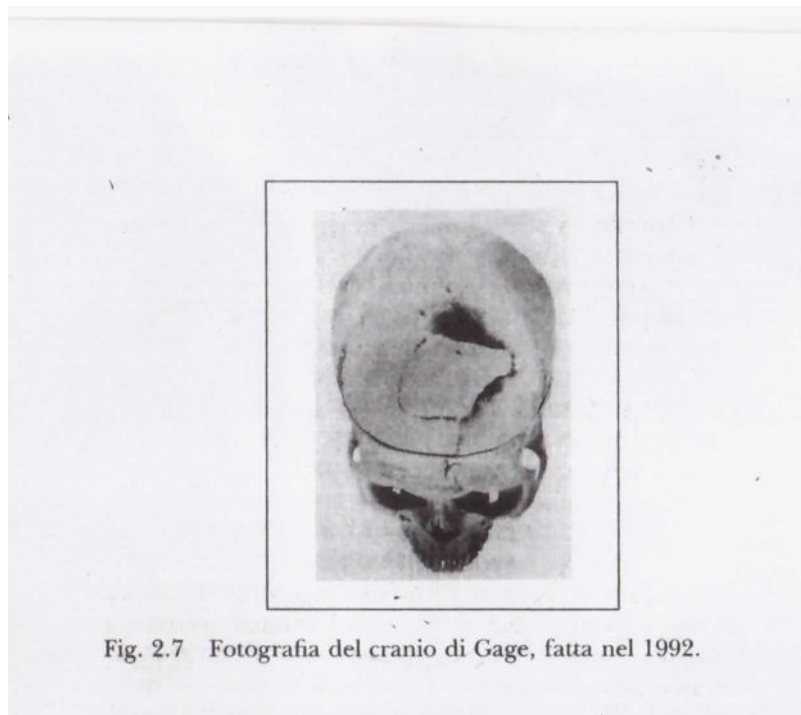


Fig. 2.8 In alto, ricostruzione del cervello e del cranio di Gage con la possibile traiettoria della barra di ferro. In basso, vista dall'interno dei due emisferi cerebrali, per mostrare come la barra lese le strutture dei lobi frontali su entrambi i lati.

L'errore di Cartesio

Non sarebbe stato possibile esporre i miei punti di vista, in questa conversazione, senza citare Cartesio, simbolo di una serie di idee sul corpo, sul cervello e sulla mente che in un modo o nell'altro continuano a influenzare la scienza e la cultura occidentali. A me provocano disagio, come si è visto, sia la concezione dualistica per la quale Cartesio scinde la mente dal cervello e dal corpo (e che, nella versione estrema, è meno dominante), sia le varianti moderne di essa: l'idea, per dirne una, che mente e cervello siano sì in relazione, ma solo nel senso che la mente è il programma (il software) che gira in un pezzo di hardware di un calcolatore chiamato cervello; oppure che il cervello e il corpo siano sì in relazione, ma solo nel senso che il primo non può sopravvivere senza il supporto vitale del secondo.

Qual era, allora, l'errore di Cartesio? O meglio, quale errore di Cartesio io intendo isolare, senza rispetto né gratitudine? Si potrebbe cominciare con una rimostranza: rimproverargli di avere convinto i biologi ad adottare (fino ai nostri giorni) meccanismi simili a orologi come modelli per i processi della vita. Ma questo forse non sarebbe proprio corretto; e allora si potrebbe continuare con il « Penso, dunque sono ». L'enunciato, il più famoso di tutta la storia della filosofia, appare per la prima volta in francese (« Je pense donc je suis ») nella parte quar-

ta del *Discours de la méthode* (1637) e poi in latino (« Cogito ergo sum ») nella parte prima dei *Principia Philosophiae* (1644).³ Preso alla lettera, esso esprime esattamente il contrario di ciò che io credo vero riguardo alle origini della mente e riguardo alla relazione tra mente e corpo; esso suggerisce che il pensare, e la consapevolezza di pensare, siano i veri substrati dell'essere. E siccome sappiamo che Cartesio immaginava il pensare come un'attività affatto separata dal corpo, esso celebra la separazione della mente, la « cosa pensante » (*res cogitans*), dal corpo non pensante, dotato di estensione e di parti meccaniche (*res extensa*).

E tuttavia, assai prima dell'alba dell'umanità gli esseri erano esseri. A un certo punto dell'evoluzione, una coscienza elementare ebbe inizio. Con essa arrivò una mente, semplice; aumentando la complessità della mente, sopravvenne la possibilità di pensare e, ancora più tardi, di usare il linguaggio per comunicare e organizzare meglio il pensiero. Per noi, allora, all'inizio vi fu l'essere e solo in seguito vi fu il pensiero; e noi adesso, quando veniamo al mondo e ci sviluppiamo, ancora cominciamo con l'essere e solo in seguito pensiamo. Noi siamo, e quindi pensiamo; e pensiamo solo nella misura in cui siamo, dal momento che il pensare è causato dalle strutture e dalle attività dell'essere.

Se proviamo a ricollocare l'enunciato cartesiano nel contesto a cui appartiene, possiamo chiederci se potrebbe avere un significato diverso da quello che ha assunto. È possibile leggerlo come il riconoscimento della superiorità del sentire e del ragionare consci, senza alcuna presa di posizione riguardo alla loro origine, sostanza o persistenza? È possibile che esso sia servito al sagace scopo di contenere pressioni religiose delle quali Cartesio era ben consapevole? Questa seconda possibilità non vi è modo di accertarla. Sulla propria tomba, Cartesio volle che fosse apposta una frase di Ovidio: « Bene qui latuit,

bene vixit » (*Tristia*, III, 4, v. 25), alla quale (sembra) egli fece ricorso di frequente. « È vissuto bene chi bene si è celato »: una criptica abiura dal dualismo, forse? Quanto alla prima, di possibilità, io credo d'altra parte che Cartesio intendesse *anche* precisamente quello che scrisse. Alla comparsa di quelle celebri parole, egli gioisce della scoperta di una proposizione così irrefutabilmente vera che non v'è scetticismo in grado di incrinarla:

« E notando che questa verità: *io penso, dunque sono*, era così solida e sicura che tutte le più stravaganti supposizioni degli scettici non erano capaci di scuoterla, giudicai di poterla accogliere senza scrupolo come il primo principio della filosofia che cercavo ». ⁴

Qui Cartesio voleva giungere a un fondamento logico per la propria filosofia, e il suo enunciato non differiva dal « Fallor ergo sum » di sant'Agostino. ⁵ Ma poche righe più avanti egli chiarisce la propria affermazione in modo che non si presta a equivoci:

« Pervenni in tal modo a conoscere che io ero una sostanza, la cui intera essenza o natura consiste nel pensare, e che per esistere non ha bisogno di alcun luogo, né dipende da alcuna cosa materiale. Di guisa che questo io, cioè l'anima, per opera della quale io sono quel che sono, è interamente distinta dal corpo, ed è anzi più facile a conoscere di questo; e anche se questo non fosse affatto, essa non cesserebbe di essere tutto quello che è ». ⁶

Eccolo, l'errore di Cartesio: ecco l'abissale separazione tra corpo e mente – tra la materia del corpo, dotata di dimensioni, mossa meccanicamente, infinitamente divisibile, da un lato, e la « stoffa » della mente, non misurabile, priva di dimensioni, non attivabile con un comando meccanico, non divisibile; ecco il suggerimento che il giudizio morale e il ragionamento e la sofferenza che viene dal dolore fisico o da turbamento emotivo possano esistere separati dal

corpo. In particolare: la separazione delle più elaborate attività della mente dalla struttura e dal funzionamento di un organismo biologico.

Ma perché – si potrebbe obiettare – far tante questioni con Cartesio e non con Platone, le cui opinioni sul corpo e sulla mente, quali si possono ritrovare nel *Fedone*, sono assai più irritanti? Perché prendersela per questo particolare errore di Cartesio – che, dopo tutto, ne fece anche altri, e più marchiani? Ad esempio, egli credette che fosse il calore a fare circolare il sangue, e che le minuscole, finissime particelle di sangue si distillassero in « spiriti animali », capaci di muovere i muscoli. Perché non criticarlo per queste sue convinzioni? La ragione è semplice: si sa da molto tempo che su questi punti egli si sbagliava, e la questione di come e perché il sangue circoli è stata risolta nel modo più soddisfacente. Non si può dire lo stesso per quanto riguarda mente, cervello e corpo, e qui l'errore di Cartesio mantiene una certa influenza: molti ancora pensano che le sue opinioni siano di per sé evidenti, e che non richiedano alcun riesame.

L'idea cartesiana di una mente scissa dal corpo può essere stata, attorno alla metà del ventesimo secolo, l'origine della metafora della mente come programma di software. Infatti, se la mente può essere separata dal corpo, forse si può tentare di comprenderla senza alcun ricorso alla neurobiologia, senza che occorra lasciarsi influenzare da conoscenze di neuroanatomia, di neurofisiologia, di neurochimica. Ed è interessante notare questo paradosso: molti scienziati cognitivisti, convinti di poter indagare la mente senza rifarsi alla neurobiologia, non si considererebbero dualisti.

Può esservi qualche venatura cartesiana di separatezza dal corpo anche dietro il pensiero di quei neuroscienziati i quali sostengono che è possibile dare piena spiegazione della mente solo in termini

di eventi cerebrali, lasciando ai margini il resto dell'organismo e l'ambiente fisico e sociale che lo circonda – e anche il fatto che parte dell'ambiente è essa stessa un prodotto delle precedenti attività dell'organismo. Io respingo questa limitazione, non perché la mente non sia correlata in via diretta con l'attività del cervello (è evidente che lo è), quanto perché la formulazione restrittiva è incompleta senza che ve ne sia necessità e insoddisfacente dal punto di vista umano: dire che la mente viene dal cervello è affermazione irrefutabile, ma io credo che sia meglio precisarla, e considerare le ragioni per le quali i neuroni del cervello si comportano in modo così meditato: è questa, a mio parere, la questione critica.

Sembra, inoltre, che l'idea di una mente distaccata dal corpo abbia foggiato il peculiare modo in cui la medicina occidentale affronta lo studio e il trattamento della malattia (si veda il « Post scriptum »). La scissione cartesiana permea sia la ricerca sia la pratica medica; con il risultato che le conseguenze psicologiche delle malattie del corpo in senso stretto (le cosiddette « vere » malattie) di solito vengono trascurate, e prese in considerazione, semmai, solo in un secondo momento. Ancora più trascurati sono i fenomeni inversi, cioè gli effetti somatici di conflitti psicologici. È suggestivo pensare che Cartesio contribuì a modificare il corso della medicina, a far sì che essa deviasse dall'orientamento organico, o meglio « organismico » (« la-mente-è-nel-corpo »), che era prevalso dai tempi di Ippocrate fino al Rinascimento. Quanto sarebbe stato infastidito da Cartesio, Aristotele, se l'avesse conosciuto!

Svariate versioni dell'errore di Cartesio celano che le radici della mente umana si trovano in un organismo biologicamente complesso ma fragile, finito e unico; tengono nell'ombra la tragedia implicita nel conoscere tale fragilità, finitezza e unicità. E se gli esseri umani non riescono a vedere l'intrinseco

dramma di un'esistenza conscia, tanto meno si sentiranno chiamati a fare qualcosa per attenuarlo, e possono avere meno rispetto per il valore della vita.

I fatti che ho esposto riguardo a sentimenti e ragione, assieme agli altri che ho discusso sulle interconnessioni tra cervello e corpo, confortano l'idea generale presentata all'inizio del libro: una piena comprensione della mente umana richiede una prospettiva integrata: la mente non solo deve muovere da un « cogito » non fisico al regno dei tessuti biologici, ma deve anche essere correlata con un organismo intero, in possesso di un cervello e di un corpo integrati e in piena interazione con un ambiente fisico e sociale.

Ma la mente davvero intrisa nel corpo per come la vedo io non abbandona i livelli più raffinati di attività, quelli che ne costituiscono l'anima e lo spirito. Nella mia prospettiva, anima e spirito, con tutta la loro dignità e misura umana, sono ora stati, complessi e unici, di un organismo. Forse la cosa davvero indispensabile che noi come esseri umani possiamo fare è ricordare a noi stessi e agli altri, ogni giorno, la nostra complessità, fragilità, finitezza e unicità. E qui sta il difficile; non nel muovere lo spirito dal suo piedistallo sul nulla a un qualche sito, preservandone dignità e importanza, ma nel ricoscerne la vulnerabilità, le umili origini, e tuttavia continuare a fare appello alla sua guida. Compito difficile davvero, ma indispensabile, rinunciando al quale sarebbe assai meglio lasciare non corretto l'errore di Cartesio.

H. A. Simon
LE SCIENZE DELL'ARTIFICIALE
CAPITOLO PRIMO

LA CONOSCENZA DEL MONDO NATURALE
E DEL MONDO ARTIFICIALE

Tre secoli circa dopo Newton, abbiamo assimilato a fondo il concetto di scienze naturali, piú esattamente di scienze biologiche e fisiche.

Scienza naturale è quell'insieme di conoscenze relative ad una classe di cose — oggetti o fenomeni — esistenti nel mondo: conoscenze circa le caratteristiche e le proprietà di cui sono dotate queste cose e il modo in cui esse si comportano ed interagiscono tra loro.

Funzione primaria delle scienze naturali è rendere semplice il meraviglioso: mostrare che la complessità, opportunamente considerata, altro non è che la maschera della semplicità; identificare un modello nascosto nel caos apparente. Uno dei primi fisici olandesi, Simon Stevin, dimostrò con un elegante disegno (figura 1) che la legge dei piani inclinati deriva «in modo autoevidente» dalla impossibilità del moto perpetuo. Infatti l'esperienza e la ragione ci dicono che la catena di palline rappresentata nella figura non può ruotare né a destra né a sinistra bensí deve restare immobile. (Dal momento che la rotazione non cambia nulla nella configurazione, se mai la catena si muovesse, dovrebbe muoversi all'infinito). Dato che la parte inferiore della catena pende in modo simmetrico, possiamo eliminarla senza turbare l'equilibrio. Ma ora le palline che si trovano sul lato lungo del piano, bilanciano quelle del lato piú corto e ripido; e i loro numeri relativi sono in rapporto inverso ai seni degli angoli con cui i piani sono inclinati.

Stevin era cosí soddisfatto della sua costruzione che la inserí in una cornice con l'iscrizione

WONDER, EN IS GHEEN WONDER

che significa «Meraviglioso, ma non incomprensibile».

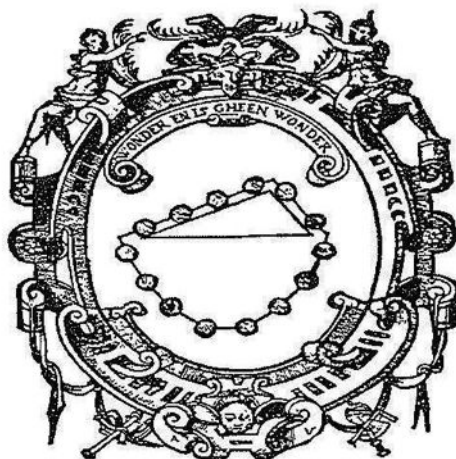


FIG. 1. Lo schizzo fatto da Simon Stevin per dimostrare la sua derivazione della legge dei piani inclinati.

Ecco il compito delle scienze naturali: rivelare che il meraviglioso non è incomprensibile, mostrare come si può capirlo senza peraltro far svanire la meraviglia. Quando infatti il meraviglioso è stato spiegato, il disegno nascosto è stato rivelato, ecco che ci stupiamo nuovamente per il modo in cui la complessità è stata costruita a partire da ciò che è semplice. L'estetica delle scienze naturali e della matematica è simile a quella della musica e della pittura: entrambe consistono nella scoperta di una struttura parzialmente celata.

Il mondo in cui viviamo oggi è sempre più un mondo artificiale, fatto dall'uomo¹, e non un mondo naturale. In quasi tutti gli elementi che ci circondano vi sono tracce dell'intervento artificiale dell'uomo. Viviamo la maggior parte delle nostre ore ad una temperatura che è tenuta artificialmente sui 21-22 gradi, all'aria che respiriamo viene aggiunta o tolta umidità, e le impurità che aspiriamo sono in gran parte prodotte (e filtrate) dall'uomo.

Inoltre, per molti di noi — quelli cioè che non fanno un lavoro manuale — la parte rilevante dell'ambiente è costituita soprattutto da una serie di artefatti, detti «simboli», che ci pervengono attraverso gli occhi e le orecchie sotto forma

di linguaggio scritto e parlato e che riversiamo nell'ambiente — come io stesso faccio in questo momento — con la bocca o le mani. Le leggi che regolano queste sequenze di simboli, le leggi che regolano le occasioni in cui noi li emettiamo e li riceviamo, le cause determinanti del loro contenuto sono tutte conseguenze del nostro artefatto collettivo.

Si potrebbe obiettare che esagero l'artificialità del nostro mondo. L'uomo deve obbedire alla legge di gravità proprio come un sasso e, come un organismo vivente, deve per la sua alimentazione e per molti altri aspetti dipendere dal mondo dei fenomeni biologici. Posso riconoscere di aver esagerato, ma di poco. Dire che un astronauta o un pilota di aereo obbediscono alla legge di gravità e che quindi sono fenomeni perfettamente naturali, corrisponde a verità; ma questa verità richiede un uso piuttosto sofisticato dell'espressione «obbedire» ad una legge naturale. Per Aristotele non era naturale che gli oggetti pesanti si sollevassero o quelli leggeri cadesse-
ro (*Fisica*, libro IV); ma probabilmente la nostra conoscenza del «naturale» è più approfondita della sua.

In modo analogo dobbiamo evitare di dare uguale significato a «biologico» e a «naturale». Una foresta può essere un fenomeno naturale; una fattoria certamente non lo è. Proprio quei generi di sostentamento come il grano e il bestiame, da cui l'uomo dipende per la sua alimentazione, sono artefatti prodotti dalla sua ingegenosità. Un campo arato non fa parte della natura né più né meno di una strada asfaltata.

Questi esempi delimitano i termini del problema: le cose che chiamiamo artefatti non sono distaccate dalla natura. Non sono autorizzate ad ignorare o violare le leggi naturali, ma nello stesso tempo sono adatte agli obiettivi e agli scopi dell'uomo. Sono ciò che sono allo scopo di soddisfare il desiderio dell'uomo di volare o di mangiare bene. Come cambiano gli obiettivi dell'uomo, così cambiano anche i suoi artefatti — e viceversa, naturalmente.

Se studiare questi oggetti e questi fenomeni in cui si incontrano fini umani e leggi naturali è una scienza, essa deve disporre di mezzi per mettere in relazione queste due disparate componenti. La natura di tali mezzi e le loro implicazioni per certe aree della conoscenza — l'economia, la psicologia e in particolare la progettazione — costituiscono l'argomento principale di questo libro.

1. *L'artificiale*

Le scienze naturali sono la conoscenza degli oggetti e dei fenomeni naturali. Ci chiediamo se non possono esserlo anche le scienze «artificiali» — cioè la conoscenza riguardante gli oggetti ed i fenomeni artificiali. Purtroppo il termine «artificiale» ha una sfumatura peggiorativa di cui bisogna liberarsi prima di andare avanti.

Nel mio dizionario, la definizione della voce «artificiale» è: «Prodotto con artificio, in opposizione a ciò che è opera di natura; non genuino o naturale; artificioso; che non riguarda l'essenza». I suoi sinonimi sono: affettato, fittizio, manufatto, falso, finto, simulato, spurio, inventato, innaturale. I contrari: vero, genuino, onesto, naturale, reale, veritiero, schietto. Sembra che il nostro linguaggio rifletta la profonda diffidenza che l'uomo nutre per i suoi prodotti. Non tenterò di analizzare la validità di tale giudizio, né di ricercare le sue possibili «radici» psicologiche. Quello che è importante capire è che io uso il termine «artificiale» nel senso più neutro possibile: per me significa solo fatto dall'uomo, in opposizione a naturale².

In certi contesti noi facciamo una distinzione tra «artificiale» e «sintetico». Ad esempio, una gemma di vetro colorato che assomiglia ad uno zaffiro viene detta artificiale, mentre una gemma prodotta dall'uomo e chimicamente non distinguibile dallo zaffiro è detta sintetica. La stessa distinzione si fa spesso tra la gomma «artificiale» e quella «sintetica». Dunque certi oggetti artificiali sono imitazioni di cose esistenti in natura e per l'imitazione si possono usare le stesse materie prime della natura o materiali del tutto diversi.

Introducendo il concetto di «sintesi» e di «artefatto», entriamo nel regno della tecnica. Infatti spesso a «sintetico» viene attribuito il significato più ampio di «progettato» o «composto». Diciamo che la tecnica ha a che fare con la «sintesi», mentre la scienza si occupa di «analisi». Gli oggetti sintetici o artificiali — e più precisamente la progettazione di oggetti artificiali con caratteristiche desiderate — costituiscono l'obiettivo principale dell'attività e delle capacità tecniche. L'ingegnere, e più in generale il progettista, si interessano di come dovrebbero essere le cose — di come dovrebbero essere per *raggiungere certi obiettivi e funzionare*. Per-

ciò l'eventuale scienza dell'artificiale dev'essere molto affine a una scienza dell'ingegneria, ma molto diversa, come vedremo nel quinto capitolo, da quella che viene correntemente indicata col nome di «ingegneria». Con i termini «scopi» e «doveri», introduciamo nel quadro anche la dicotomia tra normativo e descrittivo. Le scienze naturali hanno trovato il modo di escludere il normativo e di occuparsi solo del come stanno le cose. Possiamo o dobbiamo mantenere questa esclusione quando dai fenomeni naturali passiamo a quelli artificiali, dall'analisi alla sintesi^{3?}

A questo punto abbiamo individuato quattro criteri che distinguono l'artificiale dal naturale; possiamo quindi definire i confini delle scienze dell'artificiale:

- 1) gli oggetti artificiali sono sintetizzati dall'uomo (anche se non sempre o di solito con molto discernimento);
- 2) gli oggetti artificiali possono imitare le apparenze delle cose naturali ma, per uno o più aspetti, non posseggono la realtà di queste;
- 3) gli oggetti artificiali possono essere caratterizzati in termini di funzioni, scopi, adattamento;
- 4) specialmente nella fase della progettazione, gli oggetti artificiali vengono spesso discussi in termini prescrittivi oltre che descrittivi.

[-----]

brevetto.

ola Simon - Le Scienze dell'artificiale

4. Il calcolatore come artefatto

Nessun artefatto inventato dall'uomo si presta tanto a questo tipo di descrizione funzionale quanto il calcolatore numerico. È veramente proteiforme, perché le uniche proprietà che è possibile rilevare nel suo comportamento (quando si «comporta bene!») sono le proprietà organizzative. La velocità con la quale esegue le operazioni fondamentali ci può permettere di fare delle deduzioni circa le sue componenti fisiche e le loro leggi naturali; la velocità di trasmissione dei dati, ad esempio, ci permette di eliminare un eventuale tipo di componenti «lente». Per il resto, quasi nessuna delle affermazioni pertinenti che si possono fare su un calcolatore funzionante, ha qualche riferimento con la natura specifica dello *hardware*. Il calcolatore è un'organizzazione di componenti funzionali elementari in cui molto approssimativamente solo la

funzione svolta da tali componenti ha importanza per il comportamento dell'intero sistema¹⁵.

4.1. *I calcolatori come oggetti astratti*

Questa forte tendenza all'astrazione, propria del mondo dei calcolatori, consente di introdurre agevolmente la matematica nello studio della loro teoria — e ha indotto alla conclusione errata che, quando si costruirà una scienza dei calcolatori, sarà necessariamente una scienza matematica e non empirica. Permettetemi di prendere in esame uno alla volta questi due punti: l'importanza della matematica per i calcolatori e la possibilità di studiare i calcolatori in modo empirico.

L'argomento base di alcune importanti teorizzazioni, cui dette inizio John von Neumann, è stata l'affidabilità dei calcolatori. Il problema è: come costruire un sistema affidabile con parti non affidabili. Si noti che non si tratta di un problema di fisica o di ingegneria fisica. Si presume che i tecnici abbiano fatto del loro meglio, tuttavia le parti sono ancora inaffidabili. Solo il modo in cui si organizzano queste componenti può permetterci di superare l'inaffidabilità.

Per arrivare a formulare bene il problema, dobbiamo sapere qualcosa di più sulla natura di queste parti inaffidabili. In questo caso ci può essere d'aiuto il fatto che *qualsiasi* calcolatore può essere costruito mettendo insieme una serie di elementi fondamentali semplici. Ad esempio, possiamo prendere come elementi fondamentali i cosiddetti neuroni di Pitts-McCulloch. Come indica il loro nome, queste componenti furono dotate di caratteristiche anatomiche e funzionali analoghe a quelle che si suppone abbiano i neuroni del cervello ma con un grado di astrazione molto elevato. Sono formalmente isomorfe con le specie più semplici di circuiti di commutazione — i circuiti «and» (di congiunzione), «or» (di disgiunzione) e «not» (di negazione). Supponiamo ora di dover costruire un sistema con tali elementi e che ogni parte elementare abbia una precisa probabilità di funzionare in modo non corretto. Il problema è organizzare gli elementi e le loro interconnessioni in maniera tale che il sistema completo funzioni in modo affidabile.

Il punto fondamentale, in questo contesto, è che le parti

potrebbero essere neuroni o relè o transistor. Le leggi naturali che governano i relè sono molto ben conosciute, mentre non lo sono altrettanto quelle che governano i neuroni. Ma questo fatto non ha importanza, perché ciò che conta per questa teoria è che le componenti abbiano il livello di inaffidabilità e che siano collegate nel modo specificato.

Questo esempio dimostra che possiamo costruire la teoria matematica di un sistema o simulare tale sistema anche se non disponiamo di un'adeguata microteoria delle leggi naturali che governano le componenti del sistema. Tale microteoria può benissimo essere irrilevante.

4.2. *I calcolatori come oggetti empirici*

Passiamo ora alla possibilità di una scienza *empirica* dei calcolatori — ben distinta dalla fisica degli stati solidi o dalla fisiologia di componenti organiche¹⁶.

Nella realtà empirica, quasi tutti i calcolatori hanno in comune certi aspetti organizzativi. Possono quasi tutti essere divisi in un elaboratore attivo (il «Mulino» di Babbage) e in una memoria (il «Magazzino» di Babbage) collegati ad un'unità di entrata e uscita. (Alcuni dei sistemi più grandi, un po' come le colonie di alghe, sono costituiti da sistemi più piccoli, dotati di tutte queste componenti o solo di alcune). Potrei semplificare ulteriormente. Questi calcolatori sono tutti in grado di immagazzinare dei simboli (il programma), che possono essere interpretati da una componente destinata al controllo del programma e all'esecuzione delle relative operazioni. In quasi tutti è limitatissima la capacità di operare in modo simultaneo, parallelo; si tratta sostanzialmente di sistemi che fanno una cosa alla volta. In genere i simboli, prima che si possa interpretarli ed eseguire così le relative operazioni, devono essere trasferiti dalle componenti della memoria più grande nell'unità centrale di elaborazione. Si tratta insomma di sistemi capaci solo di azioni fondamentali semplici: registrare simboli, copiare simboli, trasferire simboli, cancellare simboli e confrontare simboli.

Dato che nel mondo queste macchine sono ora così numerose e dato che le proprietà che le descrivono risultano appartenere anche al sistema nervoso centrale dell'uomo, non c'è niente che ci impedisca di sviluppare una storia naturale

dei calcolatori. Possiamo studiarli come fossero conigli o scoiattoli e scoprire il loro comportamento di fronte ad un'ampia gamma di stimoli ambientali. Fintanto che il loro comportamento riflette le caratteristiche funzionali generali descritte e non dipende dai particolari dei materiali con cui sono fatti (*hardware*) possiamo costruire una teoria generale — ma empirica — dei calcolatori.

Le ricerche fatte per progettare i sistemi di elaborazione in *time-sharing* (si tratta della condivisione della stessa macchina da parte di piú utenti nello stesso tempo), sono un valido esempio di come è possibile studiare il comportamento di un calcolatore considerandolo un fenomeno empirico. Disponiamo solo di frammenti di teoria per orientare il progetto di un sistema di *time-sharing* o per prevedere come un sistema che realizza un dato progetto si comporterà in un ambiente di utenti che gli fanno molte richieste contemporaneamente. La maggioranza dei progetti hanno presentato all'inizio delle gravi deficienze; e la maggior parte delle previsioni sulle loro prestazioni si sono rivelate sorprendentemente inesatte.

In queste circostanze, la strada migliore da seguire per lo sviluppo e il miglioramento dei sistemi di *time-sharing* consiste nel costruirli e nello stare a vedere come poi si comportano. Ed è proprio quello che si è fatto. Sono stati costruiti, modificati e migliorati per stadi successivi. Forse la teoria avrebbe potuto anticipare gli esperimenti e non renderli necessari. Ma in realtà non è stato così e non mi risulta che nessuno di quelli che conoscono a fondo questi sistemi eccessivamente complessi abbia idee precise sul modo in cui la teoria avrebbe potuto farlo. Per capire questi sistemi, è necessario costruirli ed osservare il loro comportamento¹⁷.

Per restare nello stesso tema, anche i programmi per calcolatori studiati per mettere in atto dei giochi o trovare le dimostrazioni di teoremi matematici, operano in ambienti eccessivamente grandi e complessi. Anche quando si tratta di programmi non eccessivamente grandi e complicati (in confronto naturalmente ai sistemi operativi e di controllo dei grandi calcolatori), sappiamo troppo poco del loro ambiente di lavoro per poter prevedere con esattezza come funzioneranno e con quale grado di selettività riusciranno a cercare soluzioni di problemi.

Anche questa volta, l'analisi teorica deve avere il soste-

gno di grandi quantità di lavoro sperimentale. Un numero sempre maggiore di pubblicazioni su questi esperimenti ci dà un'idea precisa dell'efficacia con cui certi congegni euristici ci permettono di ridurre lo spazio di ricerca e quindi il numero dei tentativi che si devono fare per trovare la soluzione del problema. Nella dimostrazione di teoremi, ad esempio, la potenzialità euristica, basata e pilotata dalla ricerca empirica, ha fatto grandi passi avanti: si pensi all'uso del teorema di Herbrand, del principio della risoluzione, del principio dell'insieme degli argomenti favorevoli e così via¹⁸.

4.3. *Calcolatori e pensiero*

Ampliando e approfondendo la nostra conoscenza — teorica e empirica — dei calcolatori, scopriremo che in gran parte il loro comportamento è governato da semplici leggi naturali, e che ciò che sembrava complesso nei loro programmi era, in buona parte, complessità dell'ambiente al quale il programma cercava di adattare il suo comportamento.

Nella misura in cui questa prospettiva può diventare realtà, emerge, per la simulazione mediante calcolatore, un ruolo estremamente importante: quello di strumento per arrivare ad una conoscenza più profonda del comportamento umano. Perché se è l'organizzazione delle componenti, e non le loro proprietà fisiche, che determina in larga parte il comportamento e se i calcolatori sono organizzati in certa misura ad immagine dell'uomo, allora il calcolatore diviene evidentemente un mezzo per esplorare le conseguenze di ipotesi organizzative sul comportamento umano. La psicologia può progredire senza attendere che la neurologia risolva i problemi della «componentistica» umana — per quanto interessante e importante si riveli la natura di tali componenti.

5. *Sistemi simbolici: artefatti razionali*

Il computer fa parte di un'importante famiglia di artefatti chiamata «sistemi simbolici», o, più esplicitamente, «sistemi simbolici fisici»¹⁹. Un altro importante membro della famiglia (alcuni di noi pensano, antropomorficamente, che sia il più importante) è la mente umana e il cervello. È con

questa famiglia di artefatti, e in particolare con la sua versione umana, che avremo a che fare in questo libro. La quintessenza dei sistemi simbolici è l'artificialità, dato che l'adattamento ad un ambiente è la loro unica «ragione di esistenza». Si tratta di sistemi che elaborano informazioni (*information-processing*) e che sono volti ad uno scopo (*goal-seeking*), di solito inseriti o, meglio, incorporati in sistemi più ampi all'interno dei quali svolgono un servizio.

5.1. *Capacità di base dei sistemi simbolici*

Un sistema simbolico fisico possiede un insieme di entità, chiamate simboli. Si tratta di rappresentazioni fisiche (ad esempio segni di gesso su una lavagna) che entrano come componenti di strutture simboliche (talvolta chiamate «espressioni»). Come ho già sottolineato nel caso dei computer, un sistema simbolico possiede alcuni processi semplici che operano su strutture simboliche — processi cioè che creano, modificano, copiano e distruggono simboli²⁰. Un sistema simbolico fisico è una macchina che, muovendosi nel tempo, produce una collezione di strutture simboliche in evoluzione. Le strutture simboliche possono servire, e di fatto di solito servono, come rappresentazioni interne (ad esempio «immagini mentali») degli ambienti a cui i sistemi simbolici cercano di adattarsi. Esse permettono la costruzione di un modello dell'ambiente più o meno veridica e più o meno dettagliata e, di conseguenza, rendono possibile ragionarvi su. Naturalmente se vogliamo che tale capacità sia veramente utile per il sistema simbolico in questione, bisogna avere a disposizione «mani» e «finestre» sul mondo. Bisogna cioè avere degli strumenti che ci permettano di acquisire informazioni dall'ambiente esterno e di codificarle come simboli interni. Inoltre si deve disporre di strumenti per produrre simboli che permettano di intervenire sull'ambiente. Infine si devono poter usare i simboli per «designare» oggetti, relazioni e azioni nel mondo esterno al sistema.

I simboli possono anche designare processi interpretabili ed eseguibili dal sistema simbolico. Di conseguenza i programmi che governano il comportamento di un sistema simbolico possono venire immagazzinati, insieme ad altre strutture simboliche, nella memoria del sistema per venire eseguiti solo quando attivati.

I sistemi simbolici vengono chiamati «fisici» per ricordare al lettore che esistono come strumenti materiali, reali, fabbricati in vetro e metallo (i computer) o di carne e sangue (i cervelli). Nel passato ci siamo abituati a pensare ai sistemi simbolici della matematica e della logica come astratti e disincarnati, senza prendere in considerazione quei pezzi di carta, matite e menti umane che ci volevano per farli vivere. I computers hanno, per così dire, trasportato i sistemi simbolici dal mondo delle idee platonico al mondo empirico dei processi svolti da macchine o cervelli o da entrambi in collaborazione.

5.2. Intelligenza come calcolo

I tre capitoli che seguono si basano sull'assunto che l'intelligenza sia il lavoro di sistemi simbolici. Un po' più precisamente: l'ipotesi è che un sistema simbolico fisico del tipo appena descritto abbia i mezzi necessari e sufficienti per azioni intelligenti.

L'ipotesi è chiaramente empirica, da giudicare cioè vera o falsa sulla base dell'evidenza empirica. Uno degli scopi dei capitoli terzo e quarto sarà passare in rassegna questa evidenza empirica, che è di due tipi. Da un lato, grazie alla costruzione di programmi per computer che sono evidentemente in grado di azioni intelligenti, corroboriamo l'ipotesi della «sufficienza» (i mezzi del sistema simbolico fisico sono sufficienti a produrre intelligenza). D'altro canto, raccogliendo dati sperimentali sul pensiero umano che tendono a mostrare come il cervello umano operi come un sistema simbolico, noi corroboriamo l'ipotesi della «necessarietà», dal momento che questi dati implicano che un qualsiasi sistema intelligente da noi conosciuto (cervelli e computer) sia un sistema simbolico.

5.3. Economia: razionalità astratta

Come preludio alle nostre riflessioni sull'intelligenza umana come prodotto del lavoro di un sistema simbolico fisico, introdurremo nel corso del secondo capitolo una idealizzazione ed una astrazione «eroica» e cioè l'idealizzazione della razionalità umana racchiusa negli scrigni delle moderne teorie economiche, in particolare in quelle chiamate neoclassi-

che. Queste teorie costituiscono idealizzazioni dato che dirigono la loro attenzione principalmente all'ambiente esterno del pensiero umano, alle decisioni che sono ottimali per realizzare il sistema di scopi adattivo (e cioè la massimizzazione del profitto o dell'utilità). Tali teorie cercano di definire le decisioni sostanzialmente razionali nelle circostanze determinate dall'ambiente esterno.

Il modo in cui la teoria economica affronta e tratta i limiti della razionalità attribuibili al sistema interno — cioè alle caratteristiche del sistema simbolico fisico — tende ad essere pragmatico, e talvolta opportunistico. Nell'approccio formale all'equilibrio generale e nella cosiddetta teoria delle «attese razionali», l'eventualità che un sistema di elaborazione delle informazioni possa avere delle capacità di adattamento limitate viene per lo più ignorata. D'altro canto nel corso delle discussioni sui meccanismi di mercato e in molte teorie sulle decisioni in condizioni di incertezza, gli aspetti procedurali della razionalità vengono presi molto più sul serio.

Nel secondo capitolo esamineremo sia esempi di attenzione che esempi di cecità nei confronti dei limiti della razionalità. Dalle idealizzazioni dell'economia passeremo, nel corso dei capitoli terzo e quarto, al più sistematico studio dell'ambiente interno del pensiero — dei processi di pensiero così come di fatto si realizzano nell'ambito dei vincoli imposti dai parametri di un sistema simbolico fisico come il cervello.

NOTE AL CAPITOLO PRIMO

¹ In tutto questo libro userò «uomo» in senso neutro, comprendente entrambi i sessi, cosicché «lui», «gli», «egli» saranno dei pronomi «androgini». Non penso che si debbano abbandonare questi pronomi, ma estenderli così da riferirsi sia a donne che a uomini.

² Declino ogni responsabilità per questa scelta di termini. L'espressione «intelligenza artificiale» che mi ha indotto a questa scelta, fu coniata, mi pare, proprio al MIT. Il nostro gruppo di ricerca della RAND e della Carnegie-Mellon University ha preferito espressioni come «elaborazione complessa delle informazioni» e «simulazione di processi cognitivi». Ma in tal caso incappiamo in nuove difficoltà: secondo il dizionario, «simulare» significa «assumere o avere solo l'apparenza o la forma di qualcuno o qualcosa, senza averne la realtà; imitare; contraffare; fingere». Ad ogni modo «intelligenza artificiale» sembra un'espressione destinata a durare

ed è forse più semplice chiarirla che eliminarla. Col tempo diventerà tanto idiomatica da smettere di essere il bersaglio di una retorica a buon mercato.

³ Questo argomento sarà discusso a lungo nel quinto capitolo. Per non tenere in sospeso i lettori, posso precisare che io rimango dell'originaria idea positivista dell'irriducibilità di «dovere» a «essere», come già esposto nel capitolo III del mio *Administrative Behavior*, New York, MacMillan, 1976³; trad. it. *Il comportamento amministrativo*, Bologna, Il Mulino, 1979³. Questa posizione è assolutamente coerente con il considerare come fenomeni i sistemi naturali o artificiali rivolti a uno scopo senza prestare nessuna attenzione alle loro finalità: cfr. *ibidem*, appendice. Cfr. anche il noto articolo di A. Rosenbluth, N. Wiener e J. Bigelow, *Behavior, Purpose, and Teleology*, in «Philosophy of Science», 10 (1943).

⁴ Generalizzando l'argomentazione qui svolta sulla separabilità dell'ambiente «esterno» da quello «interno», è chiaro che dovremmo trovare questa separabilità, in grado più o meno elevato, in tutti i sistemi grandi e complessi, siano essi artificiali o naturali. Nella forma generalizzata, ciò significa che tutta la natura è organizzata in «livelli». Il capitolo settimo di questo volume, *L'architettura della complessità*, sviluppa meglio l'argomento.

⁵ La teoria dell'equivalenza funzionale dei calcolatori ha avuto un notevole sviluppo negli ultimi anni. Cfr. M.L. Minsky, *Computation: Finite and Infinite Machines*, N.J., Englewood Cliffs N.J., Prentice-Hall, 1967, capp. I-IV.

⁶ Brevetto U.S.A. 1.307.836, concesso a Arthur Simon, 24 giugno 1919.

⁷ *Ibidem*.

⁸ *Ibidem*.

⁹ Per quanto concerne il ruolo cruciale dell'adattamento o della razionalità, e dei loro limiti, in rapporto all'economia e alla teoria dell'organizzazione vedi l'introduzione alla parte IV, *Rationality and Administrative Decision Making* del mio *Models of Man*, New York, Wiley, 1957, pp. 38-41, 80-81, e di *Il comportamento amministrativo*, cit., pp. 351-355, e il cap. II di questo libro.

¹⁰ Si confronti la corrispondente affermazione circa la progettazione delle organizzazioni amministrative: «La razionalità, quindi, non determina il comportamento. Nell'ambito dell'area di razionalità il comportamento si presenta perfettamente flessibile ed adattabile alle capacità, alle conoscenze ed ai fini. Al contrario, gli elementi irrazionali e non razionali che segnano il limite della sfera di razionalità determinano il comportamento. [...] la teoria amministrativa deve portare la propria attenzione sui limiti della razionalità e sui modi con i quali l'organizzazione determina detti limiti per la persona che prende le decisioni»: *Il comportamento amministrativo*, cit., p. 352.

¹¹ Per alcune versioni del modello in letteratura cfr. A.D. Dahlberg, *National Income Visualized*, New York, Columbia University Press, 1956.

¹² A.W. Phillips, *Mechanical Models in Economic Dynamics*, in «Economics», nuova serie, 17 (1950), pp. 283-305.

¹³ Per indicare il passaggio delle scienze dal generale al particolare Simon ricorre ad un termine della psicologia dei processi cognitivi che indica una strategia dal globale al particolare (ad esempio nel riconoscimento di oggetti) o, anche, dall'alto al basso: *top-down*. [N.d.T.].

¹⁴ Questo punto è approfondito in questo volume nel capitolo VII, *L'architettura della complessità*. Più di cinquant'anni fa Bertrand Russell fece le stesse osservazioni a proposito dell'architettura della matematica. Si veda nella *Prefazione di Principia Mathematica*: «la ragione principale in favore di qualsiasi teoria dei fondamenti della matematica deve essere sempre induttiva, cioè deve trovarsi nel fatto che la teoria in questione ci consente di dedurre la matematica ordinaria. In matematica, il massimo grado di evidenza non si trova all'inizio, ma in qualche punto successivo; quindi le prime deduzioni, finché non raggiungono questo punto, danno motivo di credere alle premesse perché da esse derivano conseguenze vere, piuttosto che credere alle conseguenze perché derivano dalle premesse». La tendenza contemporanea verso i formalismi deduttivi spesso ci nasconde questo fatto fondamentale, che non è oggi men vero che nel 1910.

¹⁵ Sugli argomenti di questo capitolo e dei successivi, cfr. M.L. Minsky, *Computation: Finite and Infinite Machines*, cit.; quindi J. von Neuman, *Probabilistic Logic and the Synthesis of Reliable Organisms from Unreliable Components*, in *Automata Studies*, a cura di C.E. Shannon e J. McCarthy, Princeton, Princeton University Press, 1956.

¹⁶ A. Newell e H.A. Simon, *Computer Science as Empirical Inquiry*, in «Communications of the ACM», 19 (1976), marzo, pp. 113-126.

¹⁷ L'afflato empirico, esplorativo che accompagna la ricerca nel settore dei calcolatori è ben espresso nella relazione che Maurice V. Wilkes svolse nel 1967 durante le sue lezioni su Turing, *Computers Then and Now*, in: «Journal of the Association for Computing Machinery», 15 (1968), gennaio, pp. 1-7.

¹⁸ Si notino ad esempio i dati empirici in L. Wos, G.A. Robinson, D.F. Carson e L. Shalla, *The Concept of Demodulation in Theorem Proving*, in «Journal of the Association for Computing Machinery», 14 (1967), ottobre, pp. 698-709; e in altri articoli là citati. Cfr. anche la raccolta di programmi in *Computers and Thought*, a cura di E. Feigenbaum e J. Feldman, New York, MacGraw - Hill, 1963. In questo settore è uso comune intitolare gli articoli sui programmi euristici *Experiments with an XYZ Program*.

¹⁹ Nella letteratura viene più spesso usata la dizione *information-processing system* (sistema di elaborazione delle informazioni). Userò i due termini come sinonimi.

²⁰ Newell e Simon, *Computer Science as Empirical Inquiry*, cit., p. 116.

creare discriminazioni tra i soggetti. Questo fatto ha dato luogo a dibattiti in cui sovente argomentazioni di tipo ideologico risultano confuse con argomentazioni di tipo scientifico; al riguardo, i test di intelligenza sono stati accusati di ridotto spessore teorico. Una critica più seria rileva che non già la teoria ha ispirato la costruzione di un test, bensì la costruzione del test è proceduta «a tentoni», attraverso revisioni successive, fino a raggiungere un assetto psicometrico ritenuto soddisfacente. Perciò si tratterebbe di un'operazione tautologica, in quanto la variabile misurata sarebbe stata definita dal metodo stesso della sua misurazione (è significativa la posizione di chi ritiene che l'intelligenza sia «ciò che è misurato dai test di intelligenza»). All'opposto, si può sottolineare il ruolo che i test hanno avuto nella ricerca e nella teorizzazione psicologica. Per certi aspetti dell'intelligenza proprio l'approccio psicometrico è il più idoneo per avviare la ricerca, per fornire dati iniziali a partire dai quali impostare la riflessione, per desumere suggerimenti e spunti: alcune teorie psicologiche, p. es. quella di Sternberg, poggiano consistentemente sui risultati ricavati dalla somministrazione di test.

A. Antonietti

BIBL.: A. BINET, *Etude expérimentale de l'intelligence*, Paris 1903; A. BINET - T. SIMON, *Méthodes nouvelles pour le diagnostic du niveau intellectuelle des anormaux*, in «L'Année Psychologique», 11 (1905), pp. 245-336; L.M. TERMAN, *The Measurement of Intelligence*, Boston 1916; C. SPEARMAN, *The Abilities of Man. Their Nature and Measurement*, London 1927; C. SPEARMAN, *The Nature of Intelligence and the Principles of Cognition*, New York - London 1927²; D. WECHSLER, *The Measurement of Adult Intelligence*, Baltimora 1944; J.P. GUILFORD, *The Nature of Human Intelligence*, New York 1967; R.B. CATTELL, *Abilities*, Boston 1971; P.E. VERNON, *The Structure of Mental Abilities*, London 1971; H. GARDNER, *Frames of Mind. the Theory of multiple Intelligences*, New York 1983, tr. it. di L. Sosio, *Formae mentis: saggio sulla pluralità dell'intelligenza*, Milano 1987; R.J. STERNBERG, *Beyond I.Q. A Triarchic Theory of Human Intelligence*, Cambridge 1985, tr. it. di M. Mancini, *Teorie dell'intelligenza: una teoria tripolare dell'intelligenza umana*, Milano 1987.

► ANALISI FATTORIALE; MENTE; PERCEZIONE; PSICOMETRIA; RAGIONE; SENSAZIONE.

INTELLIGENZA ARTIFICIALE (*artificial intelligence; künstliche Intelligenz; intelligence artificielle; inteligencia artificial*). — È difficile dare

una definizione dell'intelligenza artificiale e dei suoi obiettivi che sia unanimemente condivisa dai ricercatori. L'origine della difficoltà sta anche nel fatto che da sempre l'intelligenza artificiale si è presentata sotto un duplice profilo: quello di disciplina «ingegneristica», il cui obiettivo è di costruire macchine in grado di assistere l'uomo, e magari di competere con esso, in diversi compiti soprattutto intellettuali; e quello di disciplina «psicologica», il cui obiettivo è di costruire macchine le quali, riproducendo da vicino caratteristiche essenziali dell'attività cognitiva umana, gettino una nuova luce su alcuni tradizionali enigmi della mente, ad esempio il «problema mente-corpo». Forse la definizione più comunemente accettata dell'intelligenza artificiale e dei suoi obiettivi è quella più generica: l'intelligenza artificiale si propone di riprodurre con programmi per calcolatore, o con robot mobili, comportamenti che, se osservati negli esseri umani, o più in generale negli organismi viventi, verrebbero definiti «intelligenti».

SOMMARIO: I. L'intelligenza artificiale degli esordi. - II. Passi verso l'intelligenza artificiale. - III. Altri sviluppi.

I. L'INTELLIGENZA ARTIFICIALE DEGLI ESORDI. — La nascita ufficiale dell'intelligenza artificiale risale al giugno del 1956, quando negli Stati Uniti, a Dartmouth (New Hampshire), si svolse un seminario organizzato da John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon. In quel memorabile incontro, al quale parteciparono coloro che sono ormai ritenuti i padri dell'intelligenza artificiale, vi gettarono le basi della nuova disciplina, vennero individuate alcune aree di ricerca rimaste classiche e furono discussi i lineamenti dei primi programmi per calcolatori cosiddetti intelligenti. Nei giorni di Dartmouth, pronto per girare su calcolatore era un programma ideato da Allen Newell, Clifford Shaw e Herbert Simon, il *Logic Theorist*, che dimostrava alcuni teoremi del calcolo enunciativo dei *Principia Mathematica* di Bertrand Russell e Alfred Whitehead. Ma bisogna ricordare almeno un altro programma la cui sperimentazione era cominciata qualche tempo prima, quello di Arthur Samuel per la dama, in grado di migliorare le proprie prestazioni nel corso del gioco, o apprendere, al punto di arrivare a battere ottimi giocatori. Altri programmi sarebbero presto seguiti: dalla *Geometry Machine* di Rochester e Herbert Gelert-

ner ai primi programmi per gli scacchi, stimolati da alcune precedenti intuizioni di Shannon.

Questi e altri programmi della prima intelligenza artificiale avevano una caratteristica in comune: essi erano in grado di simulare almeno i prodromi di una capacità ritenuta peculiare dell'intelligenza umana, quella di scegliere, davanti a un problema che dà luogo all'esplosione combinatoria delle mosse lecite, solo «alcune» sequenze di mosse che potrebbero portare alla soluzione. Classico è il caso del giocatore di scacchi, che ha di fronte a sé un numero ultra-astronomico di possibili mosse alternative, calcolate nell'ordine di 10^{20} . L'intelligenza fu identificata con questa capacità «selettiva» della mente umana, del resto ben documentata dalla psicologia dei processi del pensiero. Fu proprio con quest'ultima che l'intelligenza artificiale delle origini venne a confrontarsi, aspirando al ruolo di scienza della mente nelle vesti della *Information Processing Psychology*, o «psicologia dei processi dell'informazione», proposta da Newell e Simon. L'obiettivo era la costruzione di modelli computazionali, ovvero di programmi per calcolatore che simulassero i processi cognitivi umani in modo psicologicamente realistico.

II. PASSI VERSO L'INTELLIGENZA ARTIFICIALE. — Fu nel contesto di compiti come la logica e alcuni giochi e rompicapo, talvolta detti «problemi giocattolo» (*toy problem*), che sembrava dimostrarsi efficace la programmazione euristica, cioè lo studio e l'implementazione su calcolatore di strategie selettive di soluzione di problemi, o «euristiche». Le vere difficoltà cominciavano con i «problemi della vita reale» (*real-life problem*), quali comprendere una lingua, riassumere un testo, fare una diagnosi medica e così via. In questo caso, il programma deve essere in grado di gestire una quantità più o meno grande di informazioni specifiche sul compito da portare a termine, le quali influenzano la natura stessa delle euristiche. Il capostipite di questi programmi «basati sulla conoscenza» o «sistemi esperti», il *Dendral*, era specializzato nel determinare la struttura chimica di composti organici, e nacque dalla collaborazione di Edward Feigenbaum con il Nobel per la genetica Joshua Lederberg.

Indipendentemente dall'evoluzione dei sistemi esperti, che hanno conosciuto applicazioni commerciali di notevole portata, l'attenzione

per i problemi della vita reale ha contribuito a riquilibrare molte aree di ricerca che si erano mostrate intrattabili con l'approccio della prima intelligenza artificiale: dalla traduzione automatica alla comprensione del linguaggio naturale, alla visione artificiale, quest'ultima in particolare ad opera di David Marr negli anni ottanta. Alcune di queste aree sono state esplorate anche da ricercatori che continuavano a rifarsi alle ambiziose aspirazioni teoriche della prima intelligenza artificiale, le quali, a partire dalla metà degli anni settanta, confluirono in una nuova scienza della mente: la «scienza cognitiva» (H. Gardner, *The Mind's New Science: a History of the Cognitive Revolution*, New York 1987, tr. it. di L. Sosio, *La nuova scienza della mente*, Milano 1988).

Queste e altre ricerche hanno dovuto affrontare il problema di chiarire quale sistema di strutture simboliche fosse adeguato per rappresentare le conoscenze nei programmi. Ciò ha dato luogo a dispute molto vivaci all'interno della comunità dell'intelligenza artificiale, che hanno contrapposto i «logicisti» come McCarthy, sostenitori della logica quale strumento di rappresentazione della conoscenza, agli «antilogicisti», primo tra tutti Minsky. Questi proposero all'inizio degli anni settanta una struttura dati, il *frame*, nella quale fosse possibile rappresentare conoscenze sotto forma di prototipi. Altre strutture simboliche proposte in intelligenza artificiale condividono una impostazione analoga, come gli *script* di Roger Schank o i «modelli mentali» di Philip Johnson-Laird. Ma tanto per i logicisti quanto per gli antilogicisti il problema difficile è stato sempre lo stesso: quello del ragionamento, sfumato e allusivo, tipico del comune buon senso, che gli esseri umani mostrano di usare, con una duttilità ben superiore a quella di qualsiasi programma di intelligenza artificiale, quando nella vita reale si trovano davanti ad aspettative mancate o eccezioni non previste. Queste difficoltà, quella di uscire dai «micro-mondi» simulati su calcolatore per studiare le prestazioni di un programma, quella di generalizzare i risultati raggiunti da singole ricerche, rappresentano solo un inventario parziale dei problemi nei quali si è imbattuta l'intelligenza artificiale. Nel tentativo di affrontarli, la sperimentazione è proseguita con diverso grado di successo in numerosi settori. Quello dei programmi di simulazione della scoperta

scientifici, realizzati da Simon e da diversi ricercatori a partire dagli anni ottanta, è un esempio. Intorno alla metà di quegli anni Douglas Lenat ha lanciato il progetto del programma *cyc*, che dovrebbe rivelarsi in grado di gestire un volume di conoscenze sempre più ampio. Il programma *soar*, originariamente ispirato da Newell, ha proposto la sfida di una teoria unificata della cognizione. Un approccio innovativo, quello della «intelligenza artificiale distribuita», ha sviluppato un'idea che era alla base dell'architettura di *Hershey-II*, un programma risalente agli anni settanta: l'esito sono stati i «sistemi multi-agente», nei quali la ricerca della soluzione di problemi è affidata non a un unico sistema di intelligenza artificiale, ma ad una molteplicità di sistemi, o agenti, che comunicano tra di loro cooperando allo scopo finale. Altre frontiere per l'intelligenza artificiale sono l'interazione uomo-macchina e il *web*, con la costruzione di ipertesti e di motori di ricerca intelligenti, dove anche aree collaudate dell'intelligenza artificiale come la rappresentazione della conoscenza o l'elaborazione del linguaggio naturale svolgono un ruolo di rilievo.

III. ALTRI SVILUPPI. – In un libro ormai celebre pubblicato nel 1969, Minsky e Seymour Papert dimostrarono i limiti intrinseci di macchine che inizialmente si erano candidate come antagoniste ai primi programmi per calcolatore basati sulle euristiche, le cosiddette «reti neurali» artificiali, come il *Perceptron*, realizzato nel 1958 da Frank Rosenblatt nel clima della cibernetica. Questi aveva ripreso l'idea, formulata nel 1943 da Warren McCulloch e Walter Pitts, dei «neuroni formali», i quali, collegati in reti, simulassero alcune proprietà del sistema nervoso degli organismi. Il *Perceptron* di Rosenblatt era in grado di migliorare le proprie prestazioni, in compiti di discriminazione e classificazione di forme elementari, attraverso semplici algoritmi di apprendimento. Nonostante vari ricercatori, come James Anderson, Eduardo Caianiello, Teuvo Kohonen, Stephen Grossberg, abbiano continuato a lavorare un po' ovunque sulle reti neurali anche dopo la pubblicazione del libro di Minsky e Papert, la ripresa su vasta scala delle reti neurali risale ai primi anni ottanta. Tappe fondamentali di questa ripresa sono state da una parte lo sviluppo di calcolatori dotati di grandi capacità di memoria e di calcolo (le reti neurali so-

no simulate su calcolatore), dall'altra le ricerche di John Hopfield sulle memorie associative e i positivi risultati ottenuti da David Rumelhart e collaboratori mediante un algoritmo di apprendimento per correzione dell'errore noto come *backpropagation* («retropropagazione»), che permette di superare le principali limitazioni dei *Perceptron* denunciati da Minsky e Papert.

Le nuove reti neurali sono state inizialmente considerate da non pochi ricercatori come una via per uscire dalle già ricordate difficoltà dell'intelligenza artificiale e per progettare modelli cognitivi più adeguati di quelli proposti da quest'ultima, perché ritenuti più ispirati all'architettura del cervello reale. Le reti neurali posseggono una memoria distribuita, basata su procedure di tipo associativo piuttosto che a indirizzi, come nel caso del calcolatore, e rispetto all'elaborazione seriale tipica di quest'ultimo manifestano un forte parallelismo nell'elaborazione dei dati. Paul Smolensky (*On the Proper Treatment of Connectionism*, tr. it. di A. Parodi, *Il connessionismo tra simboli e neuroni*, Genova 1992) ha infine contrapposto l'approccio «subsimbolico» del connessionismo, al quale viene spesso assimilata buona parte della ricerca sulle reti neurali, all'approccio «simbolico» dell'intelligenza artificiale tradizionale.

Per certi versi l'approccio connessionista ha dato il meglio di sé integrandosi, piuttosto che contrapponendosi, all'approccio simbolico, consentendo così diverse sperimentazioni di sistemi di intelligenza artificiale «ibridi». Il connessionismo può essere visto anche come il primo passo che ha portato a mettere in discussione la plausibilità di modelli computazionali basati sulla separazione tra studio della mente e studio del cervello, che ha caratterizzato molte ricerche di intelligenza artificiale classica, soprattutto degli esordi. La cosiddetta «nuova intelligenza artificiale» si è spinta ben più oltre, spostando l'attenzione sullo stadio della percezione del mondo reale, uno stadio ritenuto il presupposto ineludibile per la spiegazione dell'attività intelligente. È l'attenzione verso questo fenomeno che ha ispirato a Rodney Brooks, nel MIT, la costruzione di semplici robot basati su architetture molto innovative, i quali riescono a muoversi coerentemente in ambienti reali senza far uso di rappresentazioni simboliche esplicite, come av-

veniva in sistemi robotici di intelligenza artificiale tradizionale. A esperienze di questo tipo, poi sfociate nella robotica cosiddetta *behavior-based*, può essere idealmente collegato un fronte di ricerche molto eterogeneo, che include la «vita artificiale», i sistemi dinamici complessi, la modellistica neurale sintetica di Gerald Edelman, la robotica evolutiva e molta robotica ispirata dalle neuroscienze e dagli studi di etologia cognitiva.

R. Cordeschi

BIBL.: E. BURATTINI - R. CORDESCHI (a cura di), *Intelligenza Artificiale*, Roma 2001; R. CORDESCHI, *The Discovery of the Artificial*, Dordrecht 2002; R. PFEIFER - C. SCHEIDER, *Understanding Intelligence*, Cambridge (Massachusetts) 1999; S. RUSSELL - P. NORVIG, *Artificial Intelligence*, Prentice Hall 2002², tr. it. di M. Aiello, N. Cancedda e B. Errico, *Intelligenza artificiale. un approccio moderno*, Torino 1998; V. SOMENZI - R. CORDESCHI (a cura di), *La filosofia degli automi*, Torino 1994².

■ CIBERNETICA; COMPUTER; CONNESSIONISMO; INFORMATICA; INTELLIGENZA; INTERNET; MACCHINE CALCOLATRICI E LOGICHE, STORIA DELLE; MENTE; MENTE, FILOSOFIA DELLA; NEUROSCIENZE; RETI NEURALI; SCIENZA COGNITIVA; VIRTUALE.

INTELLIGENZE ASTRALI (*astral intelligences*; *Gestirntintelligenzen*; *intelligences astrales*; *intelligencias astrales*). – È nella tradizione più antica dei popoli orientali – accolta poi in Grecia e negli stati ellenistici – che le stelle abbiano un carattere divino e siano governate da intelligenze celesti, volte a influire sulla regione terrestre.

Platone parla di «una specie celeste di dei (*οὐράνιον θεῶν γένος*) in grandissima parte di fuoco, per essere splendidissima e molto bella a vedere» (*Tim.*, 39-40 a). Nel *De mundo* (2, 391 b 17), attribuito ad Aristotele, si dice che «il cielo è pieno di corpi divini che chiamiamo astri»; e Aristotele (*Metaph.*, XII, 8) afferma che i movimenti delle sfere celesti non si possono spiegare se, oltre al primo motore immobile, non si ammette, per ciascuna di esse, un motore particolare relativamente indipendente. Tali motori, associati a un corpo di natura incorruttibile dal solo vincolo della mozione, sono «intelligenze», cioè enti spirituali dotati d'intelletto e volontà. Essi sono ordinati gerarchicamente, così come gerarchica e procedente dall'uno sull'altro è la loro mozione e il moto dei cieli, che si trasmette, e si traduce infine nella partecipazione causale della generazione e corruzione e di tutte le mutazioni in atto nel

cosmo. Le intelligenze astrali vengono così a costituire un elemento essenziale dell'ordine cosmico aristotelico (cfr. *Metaph.*, 1073 a-b; *De caelo*, I, 9). Plotino, sebbene sollevi gravi obiezioni contro questa teoria aristotelica, tuttavia conferma e accentua la tesi dell'animazione celeste: il sole è un dio, perché animato, e così pure gli astri (*Enn.*, V, 1). La concezione aristotelica passa nei filosofi arabi e in Tommaso (cfr. *Sum. theol.*, I, q. 70, art. 3; *C. Gent.*, II, 70), inserita nel dato di rivelazione biblica degli angeli, spiriti puri, e nella speculazione condotta sulle gerarchie angeliche; e, del resto, domina la cosmografia medievale. Alla concezione delle intelligenze astrali è connessa l'astrologia, il cui sforzo originario è di rintracciare il significato concreto dell'influsso dei cieli sul divenire. Cfr. G. Reale, *Storia della filosofia greca e romana*, vol. IX, Milano 2004, pp. 112-113.

A.M. Moschetti

■ ASTROLOGIA.

INTELLIGIBILE (gr. νοητός; lat. *intelligibilis* - *intelligible*; *intelligibel*, *verständlich*, *intelligible*, *intelligible*). – Ciò che è conoscibile soltanto mediante l'intelletto o mediante un'intuizione intellettuale, in contrapposizione a «sensibile», ossia a ciò che è dato all'intuizione dei sensi.

I primi a opporre l'intelligibile come mondo della verità al sensibile come mondo dell'apparenza e dell'opinione furono gli eleati. Platone dal canto suo affermò che le idee, costituendo la struttura del reale, sono il principio di intelligibilità del mondo sensibile. La distinzione fra sensibile e intelligibile si traduce così in una distinzione fra i gradi di perfezione dell'essere, intesi come strettamente correlati ai gradi della conoscenza di cui l'uomo è capace: a livello dei sensi o dell'opinione (*δόξα*) i gradi dell'immaginazione e della credenza; a livello della scienza (*ἐπιστήμη*) la conoscenza discorsiva e l'intellezione. In Aristotele invece il significato ontologico della distinzione si attenua perché nella sua prospettiva intelligibili sono non solo gli esseri sopransensibili, ma anche le cose stesse del mondo corporeo in quanto costituite, oltre che dalla materia, da forme sostanziali e accidentali, per loro natura universali e dunque conoscibili per astrazione dall'intelligenza stessa.

Il predetto significato ontologico torna ad accentuarsi nella tradizione neoplatonica. Filone di Alessandria introduce l'espressione «mondo intelligibile» (*κόσμος νοητός*) in riferimen-