

Appunti di intelligenza artificiale

Guida ad una antica questione e ad una nuova disciplina.

«Quindici centimetri circa di lunghezza» commentò. «Una forma rossa convoluta con un'appendice lineare verde».

«Sì» dissi incoraggiante. «Che cosa pensa che sia?»

«Difficile dirlo». Sembrava perplesso. «Non ha la simmetria semplice dei solidi platonici, pur potendo forse avere una simmetria superiore tutta sua... Potrebbe essere un'infiorescenza o un fiore, credo».

«Potrebbe essere?» sondai.

«Sì, potrebbe» confermò.

«L'annusi» suggerii, e di nuovo egli apparve un po' sconcertato, come se gli avessi chiesto di annusare una simmetria superiore. Ma accondiscese gentilmente. Improvvisamente si animò.

«Che meraviglia!» esclamò. «Una rosa precoce».

(Oliver Sacks: «L'uomo che scambiò sua moglie per un cappello», Adelphi 1986, 32)

La rapida storia dell'intelligenza artificiale che abbiamo delineato nelle puntate precedenti è apparsa così punteggiata di polemiche e di differenze prospettive che sembrerebbe non richiedere una trattazione specifica delle discussioni che la disciplina ha suscitato e tutt'ora suscita. Vi sono però almeno due buoni motivi per citare con migliore attenzione le posizioni più battaglierie. Il primo è che glissare sul tema sottrarrebbe un sottile piacere tanto a chi scrive quanto a chi legge di Ia. Spesso ci si trova a scorrere un testo sull'intelligenza artificiale solo per ritrovare qualche riferimento alla querelle pro o contro, per rivedere le formazioni in campo e magari per rinfocolare lo spirito di parte. Leggere di Ia è un po' come leggere il foglio sportivo locale prima del derby stracciatino, dove è impensabile trovar scritto: «Per i commenti, ci risentiamo dopo la partita». In secondo luogo, è un errore ritenere che, dal momento che fin dall'inizio si è molto discusso, oggi tutto sia chiaro o non abbia comunque alcun interesse. Il dibattito indica se non altro l'esistenza di punti critici da non perdere di vista, anche se, per paura di apparire provinciali, troppo spesso si preferisce snobbarli e gettarsi ansiosamente nel più rassicurante lavoro tecnico.

La discussione

Ancora Turing

Nell'articolo in cui propone di sostituire la domanda se «le macchine possono pensare» col suo pragmatico test colloquiale, Alan Turing riconosce di non avere argomenti certi per prevedere se e quando il test sarà mai superato, e si limita a esprimere la sua opinione; d'altra parte, «purché venga chiaramente messo in evidenza quali siano i fatti provati e quali siano le congetture, [...] le congetture sono di importanza fondamentale, dato che suggeriscono utili linee di ricerca». È bene però sgombrare il campo dalle obiezioni più frequenti all'idea stessa di un pensiero artificiale, chiarendo perché non si può dare per certo che l'Ia sia impossibile. Per l'asimmetria che c'è tra possibile e impossibile, non è necessario dimostrare che il progetto è realizzabile: la possibilità è qui un fatto empirico, si può soltanto tentare e «sperare — come dice Turing — che le macchine saranno alla fine in grado di competere con gli uomini in tutti i campi puramente intellettuali». La impossibilità va invece dimostrata a priori, e se le dimostrazioni non reggono, nulla autorizza a pensare che il tentativo debba per forza fallire.

Turing classifica le argomentazioni

contrarie in gruppi distinti, secondo uno stile poi ampiamente ripreso. Alcune sono riassunte con acume, come «l'obiezione della 'testa nella sabbia': le conseguenze delle macchine pensanti sarebbero terribili, speriamo che esse non possano esistere»; o ancora, come le argomentazioni «fondate su incapacità varie». Queste argomentazioni prendono la forma: «Vi concedo che possiate far fare alle macchine tutto quello cui avete accennato, ma non potrete mai costruirne una capace di fare X». A distanza di qualche decennio, questa seconda appare l'obiezione più abusata, soprattutto per la facilità con cui si sostituisce ad «X» qualsiasi cosa le macchine non abbiano ancora fatto, anche quando qualche X proposto in precedenza venga inaspettatamente conseguito. Rimandiamo al testo originale («Macchine calcolatrici e intelligenza», 1950, in *La filosofia degli automi*, a cura di Somenzi e Cordeschi, Boringhieri 1986) per le misurate controargomentazioni avanzate da Turing su queste come su altre obiezioni (l'obiezione teologica, l'argomento dell'autocoscienza, l'obiezione di Lady Lovelace, l'argomentazione fondata sulla continuità del sistema nervoso, quella sulla percezione extrasensoriale e quella del comportamento senza regole rigide). Per la particolare competenza di Turing sull'argomento, non si può non dedicare qualche parola in più alla sua opinione sulla «obiezione matematica».



I limiti della macchina

«Esiste una serie di risultati della logica matematica che possono essere usati per dimostrare che esistono delle limitazioni ai poteri delle macchine a stati discreti. Il più conosciuto di questi risultati è noto come teorema di Gödel (1931), e dimostra che in qualsiasi sistema logico sufficientemente potente possono essere formulati degli enunciati che non possono essere né provati né confutati all'interno del sistema, a meno che il sistema stesso non sia autocontraddittorio». A un risultato analogo era pervenuto Turing stesso (1936) a proposito della sua macchina digitale «a capacità infinita» (macchina di Turing universale; si veda il riquadro alla prima puntata). «Se essa è costretta a dare risposte a domande come nel gioco dell'imitazione, ci saranno alcune domande alle quali essa o darà una risposta errata, o non darà affatto risposta, quale che sia il tempo concesso per rispondere». In particolare, se la domanda riguarda il comportamento di una seconda macchina che ha certe analogie con quella interrogata (vale a dire, se la domanda coinvolge in certi modi la struttura della macchina stessa) «si può dimostrare che la risposta sarà errata o non sarà data affatto. Questo è il risultato matematico; si sostiene che esso dimostra un'incapacità della macchina alla quale l'intelletto umano non è soggetto».

La prima osservazione è che la tesi dà per scontato che la mente biologica non abbia limiti analoghi, quando parrebbe più onesto sostenere che non è neanche in grado di percepirla (solo nella persona di qualche studioso di logica è in grado di capirla...). In sostanza, l'obiezione matematica coincide con l'idea che una mente «meccanica» è condannata a non barare, a costo di dover ammettere l'errore e l'autocontraddizione, mentre l'uomo... può sempre pensare ad altro. Parrebbe insomma che proprio la «perfezione» e la superiorità logica dell'artificiale siano prese a motivo dell'impossibilità a simulare la mente umana. (La mente artificiale sarebbe dunque qualcosa di simile a Dio, il quale pur nella sua onnipotenza non può tuttavia mai errare, e

men che meno andare contro le leggi della logica. In effetti San Tommaso sarebbe oggi costretto a considerare il teorema di Gödel la prova clamorosa dell'inesistenza di Dio). Quest'idea dell'intelligenza artificiale come scienza (illusoria) della logica assoluta è tuttora molto diffusa tra i non addetti ai lavori. Autorevoli scienziati o matematici coinvolti per accidente, magari a un pranzo accademico, nella discussione sull'Ia, non esitano a sostenere che l'intelligenza artificiale è impossibile per assurdo, perché se esistesse e un computer intelligente rispondesse «sì» a una certa domanda mentre un altro computer ribattesse «no» (come appunto avviene normalmente nelle università, e certo tra persone intelligenti) crollerebbe la logica, la matematica e l'insegnamento superiore, con conseguenze a cui è meglio non pensare. Naturalmente l'Ia non ha niente a che vedere con una logica assoluta, e i programmi di intelligenza artificiale, quando già non sono pieni di errori per conto loro, fanno ogni sforzo per includere la logica «fuzzy» (o «sfumata»), la logica a più valori, le tecniche a blackboard (sulla lavagna è comodo scrivere e ancor più facile cancellare) e quant'altro permette di correggersi e di tornare sui propri passi. Alle strette, certe variazioni di «opinioni» sono espressamente generate attraverso scelte casuali — e non è necessario ricorrere a Freud per sostenere che la scelta di un numero «a caso» è più vincolata e sovradeterminata (dunque meno libera) per qualsiasi mortale che non per una routine «pseudocasuale».

Due pesi e due misure

Gran parte delle obiezioni alla possibilità di ottenere una mente artificiale non sono altrettanto tecniche. Passate di moda le ragioni teologiche o quelle fondate sulla percezione extrasensoriale, la palma del critico più radicale è stata raccolta da Hubert Dreyfus, un filosofo che ha cercato di riassumere gli altri motivi in una strategia a tutto campo. Il percorso seguito da Dreyfus si può così schematizzare: 1) l'intelligenza richiede la coscienza; 2) le mac-

chine non sono coscienti, ma seguono ciecamente degli algoritmi; 3) al contrario, l'uomo non segue algoritmi; 4) l'esperto umano, se forzato a farlo, può tentare di indicare le regole della sua condotta professionale, ma non in modo esaustivo, perché si tratta solo di una inessenziale verbalizzazione di processi più profondi di cui non ha piena coscienza; 5) l'intelligenza vera si basa sull'intuizione, sulla perspicacia, sulla capacità di cogliere sfumature e viste d'insieme: tali capacità non possono essere giustificate o comprese ulteriormente, ma restano misteriose e inaccessibili alla conoscenza. Dreyfus percorre ossessivamente lo spazio tra i due opposti estremi: l'intelligenza (per la macchina) dovrebbe essere cosciente, e non lo è; l'intelligenza (nell'uomo) è basata su capacità inconscie che non possono essere verbalizzate e quindi copiate da una macchina. In breve, la macchina non è intelligente perché non sa quel che fa; viceversa, l'uomo fa qualcosa di intelligente soprattutto quando non sa come lo sta facendo. Si tratta naturalmente di due diverse accezioni del concetto di coscienza, ma è indicativo che non se ne faccia mai una distinzione chiara ed esplicita. Sul primo punto, Dreyfus non aggiunge niente di originale e lascia volentieri il campo al collega John Searle, che vedremo più avanti; il secondo assunto assume invece una forma così radicale da vuotarsi di ogni forza dimostrativa specifica. Per Dreyfus infatti qualsiasi ragione e spiegazione di una decisione «intelligente» non è mai sufficiente a darne pienamente conto: prova ne sia che in tal caso anche una macchina potrebbe utilizzarla come regola efficace. Dreyfus giustifica questo suo modo di procedere etichettandolo come «stile polemico», dovuto al fatto che anche i suoi bersagli, i sostenitori dell'Ia, fanno ricorso più a illusioni, speranze ed equivoci verbali che a resoconti scientifici e ragionamenti deduttivi. C'è in questo una parte di verità, e l'unico effetto certo della sua battaglia è stata una maggior cautela dei ricercatori nell'extrapolare previsioni e conclusioni dai risultati parziali dei propri lavori. Alcuni degli esempi puntigliosamente rac-

colti da Dreyfus testimoniano abbagli clamorosi o promesse puntualmente dimenticate al loro scadere; tuttavia, se prendiamo sul serio il suo invito a non giudicare i risultati di una ricerca dall'entusiasmo del suo autore, non possiamo nemmeno fondare le obiezioni sull'eventuale fallimento delle promesse o sul cambio di umore del ricercatore (fenomeno, quest'ultimo, che Dreyfus si rammarica di non ritrovare con la frequenza attesa). Trattandosi dopo tutto di elementi caratteriali, non stupisce che lo stesso Dreyfus non ne vada immune: dal '64 ad oggi, egli ha via via scommesso che un computer esperto di scacchi non avrebbe mai battuto un bambino di 10 anni / un dilettante / un giocatore medio / un maestro di scacchi / il campione del mondo. Ora tende a sostenere, più ragionevolmente, che se ciò accade è in virtù della forza bruta del calcolo, dell'immensa capacità combinatoria della macchina elettronica, e non di una vera strategia intelligente; non risulta però che abbia mai studiato a fondo le strategie dei programmi di gioco esistenti. Non resta che soprassedere su tutto l'armamentario polemico di Dreyfus, nonostante l'interesse documentario degli aneddoti, delle risposte e delle controargomentazioni accumulate nelle diverse edizioni della sua opera più affascinante, *Che cosa non possono fare i computer* (1967, '72, '79, ed. it. Armando 1988), per mettere a fuoco l'idea di fondo che lo sostiene.

Contro la regola

Essa si può riassumere nella tesi che l'intelligenza non consiste nell'applicare, consciamente o inconsciamente, regole logiche o di comportamento; per converso, la fiducia che l'individuazione e l'applicazione di regole e leggi conduca a strategie intelligenti sarebbe l'abbaglio secolare in cui è caduto il riduzionismo e il razionalismo della tradizione culturale occidentale. Questa dilatazione prospettica porta a paragonare l'opera dell'ingegnere della conoscenza, il quale cerca di «elicitare» dall'esperto umano le regole di azione da trasferire sul sistema esperto, nienteme-

no che alla «maieutica» di Socrate, che va a interrogare gli uomini «più saggi» tra i suoi concittadini. Per esempio, Socrate chiede ad Eutifrone, un profeta religioso, quali siano i criteri per riconoscere la pietas: «E allora insegnami bene questa idea in sé quale è; affinché io, [...] servendome come di un modello, quell'azione che gli assomigli, di quante o tu o altri possiate compiere, questa io dica che è santa; quella che non gli assomigli, dica che non è». Sembra che Socrate voglia una regola rigorosa che possa essere usata anche dai non esperti. «Socrate condivide

questo assunto — che gli esperti di un'arte conoscono principi di azione e ragione e che ciò che conoscono devono anche saperlo dire — con i moderni ingegneri della conoscenza» («Si può accusare Socrate di cognitivismo?» in *Nuova Civiltà delle Macchine*, 1988 n. 1/2, pagg. 68-69). La conclusione di Dreyfus è paradossale: «La fenomenologia dell'acquisizione di abilità professionale ci mette in grado di capire perché gli ingegneri della conoscenza, da Socrate e Platone fino a Samuel e Feigenbaum, abbiano trovato tante difficoltà a far giungere l'esperto alla for-

L'la è pericolosa, parola di pentito

Nella «patristica» dell'intelligenza artificiale, Joseph Weizenbaum occupa un posto di tutto rilievo. Il programma Eliza, diffuso da Weizenbaum nel '66, era in grado di sostenere una «conversazione» in inglese su vari «copioni» predisposti dall'autore; la versione più nota permetteva di «parodiare il comportamento di uno psicoterapeuta non direttivo in un colloquio psichiatrico preliminare». L'impressione di parlare effettivamente con una persona era così forte da generare buffi equivoci; ma anche chi sapeva di avere di fronte un semplice programma finiva a volte per cadere nell'illusione e si assoggettava ad una vera e propria seduta psicanalitica. «Costoro spesso chiedevano di poter conversare con il sistema in privato, e dopo averlo fatto insistevano, nonostante le mie spiegazioni, nella convinzione che la macchina veramente li capisse» (Il potere del computer e la ragione umana, Edizioni Abele 1987, pag. 176).

È proprio dallo shock per la reazione entusiastica del pubblico che prende le mosse la riflessione di Weizenbaum, e più esattamente da tre considerazioni: «1) Un buon numero di psichiatri credette seriamente che il programma Doctor [derivato da Eliza] avrebbe potuto sfociare in una nuova forma, quasi completamente automatizzata, di terapia. [...] 2) Rimasi allibito nel vedere quanto rapidamente e profondamente le persone che conversavano con Doctor si lasciassero coinvolgere emotivamente dal computer, e come questo assumesse evidenti caratteri antropomorfi. [...] 3) Un'altra reazione generale, per me sorprendente, al programma Eliza fu la diffusione della

credenza che esso costituisse la soluzione generale al problema della comprensione del linguaggio naturale da parte del computer» (pagg. 24-26).

Questa improvvisa rivelazione induce Weizenbaum, autorevole professore presso il Massachusetts Institute of Technology, ad un coraggioso percorso intellettuale, che sfocia in una critica profonda alla delega delle fondamentali scelte economiche e sociali agli strumenti tecnologici forniti dall'informatica; una critica, in particolare, alle ricerche volte a meccanizzare le funzioni superiori dell'uomo, che spesso finiscono per contrabbandare come quintessenza dell'intelligenza umana semplici versioni schematiche e impoverite. Weizenbaum affronta «il potere del computer» attraverso l'esame rigoroso dei suoi fondamenti logici e tecnici: nell'opera citata, l'elegante capitolo che descrive i meccanismi di una macchina di Turing («Come funzionano i computer») è sia un prezioso manuale didattico, sia un'occasione per osservazioni epistemologiche e psicologiche molto suggestive. Agli occhi del profano, la programmazione a computer può sembrare una attività normativa fortemente consapevole: il programmatore dice quello che vuole, il computer obbedisce alle istruzioni. In realtà, nessun programmatore «sa» esattamente quello che vuole, né è sicuro di riuscire a «dirlo» al computer. Come la scrittura ordinaria, «la programmazione è piuttosto un test di comprensione» (pag. 110); ma mentre l'estrema flessibilità del linguaggio ordinario può nascondere eventuali difetti di logica, «un computer è immune dall'influenza seducente della pura eloquenza. [...] Un computer è un critico spietato» (pag. 111). «L'altra faccia della medaglia è la credenza che non si possa programmare qualcosa a meno di capirla completamente. Qui si

mulazione delle regole che pure usa, e perché le regole duramente ricavate, una volta programmate, non producono mai vera competenza. Semplicemente, l'esperto non sta seguendo alcuna regola!» (pag. 71).

Ora, si possono certamente assumere posizioni filosofiche scettiche o esistenziali per le quali nulla si può veramente spiegare e ogni distinzione concettuale resta un vano tentativo di afferrare l'inafferrabile. Di conseguenza, ogni distinzione tra conoscenza pratica e conoscenza teorica, tra saper fare e saper spiegare, tra credenza e scienza, finisce

per essere irrilevante. In questo caso però dovremmo rinunciare a parlare di intelligenza, e di intelligenza artificiale per giunta; diventa assai difficile enunciare solamente quale sia il problema che discutiamo. Se però riteniamo importante parlare di intelligenza, distinto da ciò che intelligenza non è, di conoscenza scientifica e di formulazioni esplicite di ciò che sappiamo o non sappiamo, dobbiamo per forza interessarci alla conoscenza teorica, esplicita in leggi, in regole, in regolarità o in qualsiasi modalità espressiva interpretabile in modo preciso; essa è qualcosa

di diverso dall'abilità dell'artigiano, dalla destrezza dell'equilibrista e dall'occhio magnetico del parapsicologo. L'errore ascrivibile alla «tradizione razionalistica occidentale» è se mai un errore assiologico, di valore, quando essa ritenesse comunque il procedere razionale, l'intelligenza, come il massimo o l'unico valore dell'esperienza umana. Ma non possiamo imputarle come colpa la fiducia che, quando si tratta di intelligenza, sia in ballo appunto la razionalità, la capacità di spiegare ciò che si vuole e ciò che si fa, il richiamarsi a principi regolativi.

dimentica una verità, cioè che il programmare (come, di nuovo, ogni forma di scrittura) è spesso una tecnica sperimentale. Si programma, proprio come si scrive, non perché si capisce, bensì per arrivare a capire. La programmazione è un'attività di progetto» (pag. 111).

La riflessione di Weizenbaum si allarga ad abbracciare scienza e tecnologia, e la loro traduzione in potenza e controllo. «Il potere del computer è soltanto una versione estrema di un potere che è intrinseco a tutti i sistemi che si autoconvalidano» (pag. 128). La critica all'intelligenza artificiale — o almeno alla interpretazione un po' esaltata che se ne dava negli anni '70 in America, in particolare da parte di Newell e Simon — è anche una protesta contro una visione riduttiva dell'uomo e delle considerazioni sociali, politiche e morali che invece devono sorreggere ogni intervento sui «sistemi» concreti della organizzazione sociale. Da un lato Weizenbaum non condivide l'ottimismo di chi giudica imminente la riproduzione artificiale delle capacità percettive e decisionali dell'uomo, proprio perché fondato su un riduzionismo eccessivamente semplicistico: la realtà è ben più articolata e ancora largamente sconosciuta. In secondo luogo, giudica moralmente inopportuna ogni delega di responsabilità a meccanismi tecnologici, a procedure «scientifiche», ad automatismi falsamente perfetti. È appunto la miscela esplosiva dei due atteggiamenti — riduzionismo e delega morale — che rende poco credibile ogni tentativo di supplire al difetto tecnologico con nuova più sofisticata tecnologia, e finisce per denunciare l'astrattezza della concezione della scienza come indipendente e neutrale, in linea di principio, rispetto alle sue applicazioni pratiche. Di fatto la pervasività di sistemi di pensiero autoconva-

lidanti, quali l'automatizzazione del processo decisionale mediante l'intelligenza artificiale, sancisce una fusione inestricabile tra teoria scientifica e alterazione immediata della realtà in conformità alle applicazioni della teoria stessa. Questo orwelliano riscrivere la storia — la storia attuale — a fronte di ogni teorizzazione che si presenti anche come visione ideologica del mondo, è tanto più inquietante quanto più i finanziamenti e gli indirizzi della ricerca sono monopolio dei centri di potere politico e militare. Cosa serve oggi, suggerisce Weizenbaum, investire tante risorse in progetti per il riconoscimento automatico della voce umana? Veramente i dottori ne hanno bisogno «per registrare le loro osservazioni mediche», come a volte viene risposto? Piuttosto, è certo che «queste macchine ascoltatrici, se potessero essere realizzate, renderebbero la sorveglianza delle comunicazioni verbali molto più facile di quanto sia oggi» (pag. 241), permettendo per esempio una illimitata estensione delle intercettazioni telefoniche.

Qual è in definitiva la proposta di Weizenbaum? Nei casi estremi, la risposta non può essere che il rifiuto radicale dello «strumento» tecnologico, così come si respinge la droga, la violenza, la guerra. L'introduzione all'edizione inglese del 1984 descrive per esempio l'ebbrezza e l'assuefazione all'automatismo irriflesso indotto dai videogiochi. Gli scenari artificiali dei videogames non solo si prestano evidentemente a veicolare ideologie aggressive e guerresche, ma soprattutto cancellano la capacità di riflettere sulle proprie azioni. L'intorpidimento psichico, la cesura tra la semplice prontezza di riflessi e ogni autentica intenzionalità, costituiscono un pericolo per la nostra civiltà. Rispetto all'ia, le osservazioni di Weizenbaum si propongono di temperare

l'eccessivo scientismo che negli anni '70 (il suo libro risale al 1976) aveva contagiato l'ambiente statunitense. L'informatica ha infranto le barriere fra le diverse discipline scientifiche e le stesse decisioni strategiche: se l'intelligenza artificiale si offre ormai come un possibile supporto per la gestione dei sistemi più complessi, ciò non deve comportare una abdicazione definitiva ad ogni responsabilità umana e personale, ma se mai una condivisione collettiva, a partire proprio dagli esperti del computer.

In questo senso, non è contraddittorio che Weizenbaum continui a insegnare tecnica dei computer agli studenti del Mit. Occorre che tutti conosciamo gli strumenti più complessi ed inquietanti, gli aspetti della tecnologia che più influenzano la nostra vita di ogni giorno; ed è altrettanto utile che i tecnici e i ricercatori di punta non smarriscano dietro la seduzione ipnotica della iperspecializzazione la loro intera coscienza di uomini responsabili.



Per non cadere dalla bicicletta

In parole povere, la capacità di andare in bicicletta non è una conoscenza teorica e non è una prestazione propriamente «intelligente»; altra cosa è il tentativo di darne la giustificazione, per esempio in termini di principi di statica e di dinamica fisica (per quanto la loro conoscenza astratta non basti ad evitare di cadere!). Non è difficile progettare una macchina che va in bicicletta (prescindendo dagli ostacoli sul percorso...). Un computer che calcolasse le correzioni da imprimere al manubrio in funzione della velocità, della traiettoria, della deviazione dalla verticale ecc. potrebbe forse cavarsela; è probabile però che un semplice servomeccanismo basato su un sistema di pendoli possa comportarsi ancora meglio e con più «naturalità». Tuttavia il servomeccanismo ci appare più lontano dal procedere intelligente che non il programma basato su regole: il livello sub-simbolico (o a-simbolico) non può inglobare alcuna conoscenza teorica. Alla fin fine, la distinzione tra ciò che è intelligente e ciò che intelligente non è non si identifica con la differenza tra ciò che funziona più o meno bene. La via «intelligente» è spesso un escamotage, una strategia adottata da una macchina universale (cervello, computer od altro) per un compito non costituzionalmente preordinato; non è necessariamente superiore nei risultati (il bilancio energetico di un pesce o di un uccello è infinitamente migliore di quello di un sottomarino o di un aereo, anche se i primi non hanno richiesto alcun dispiego di intelligenza).

La assoluta convinzione di Dreyfus che i sistemi esperti non possano mai eguagliare le prestazioni del professionista in carne ed ossa, fondata sull'assunto che quest'ultimo non segua regole esplicite e non abbia bisogno di conoscenze teoriche esprimibili in un qualche linguaggio simbolico, coincide con l'idea che la conoscenza scientifica non possa dare conto della tecnica e della pratica umana, e forse non sia nemmeno ad esse necessaria. È il principiante, secondo Dreyfus, che ha bisogno di teoria e di scienza: l'esperto ormai ne fa a meno, procede senza stampelle e

sostegni scientifici. Questa opinione bizzarra può forse essere sostenuta nel momento in cui consideriamo una attività umana in quanto fatto naturale, da indagare e da studiare mediante strumenti scientifici: in questo caso la conoscenza che possiamo avere dell'oggetto di indagine è parziale e astratta, perché la ricerca è un processo sempre incompleto, in cui non possiamo mai esser certi di aver raggiunto la verità, il completo adeguamento della teoria alla cosa («adeguatio intellectus rei»). È discutibile che ciò valga anche per quelle attività tecniche (e oggi sono la maggior parte) fortemente permeate di teoria e in stretto rapporto con la ricerca scientifica. Ma il punto essenziale è che se siamo interessati ai processi intelligenti e alla componente razionale dell'attività umana, non possiamo non privilegiare la componente teorica e simbolica, che deve poter essere esprimibile e comunicabile con il minimo di ambiguità ed il massimo di comprensibilità intersoggettiva. Un sistema esperto in grado di mantenere in equilibrio un robot-ciclista si dimostra più «intelligente» (ed è comunque più rilevante per l'Ia) di un acrobata da circo che non sapesse ricostruire e spiegare i suoi movimenti in sella al velocipede.

La differenza tra i procedimenti intelligenti dell'esperto umano e le sue emulazioni artificiali non risiede tanto nel fatto di seguire o meno delle regole, ma nel modo in cui l'uomo applica le regole, le interpreta e ne crea di nuove. In sostanza, si torna al tema dell'auto-

coscienza, alla distanza tra un pensiero consapevole di ragionare e un ragionamento «meccanico», «cieco», non cosciente. Qui Dreyfus non aggiunge nulla di originale e cede il passo al compagno di cordata John Searle.

Stanze cinesi

Professore di filosofia a Berkeley, noto per gli studi sul linguaggio e per la sua insistenza sulla natura irriducibilmente intenzionale degli «atti linguistici», Searle irrompe nel dibattito sull'Ia con un breve saggio del 1980, *Menti, cervelli e programmi* (trad. it. Clup Clued, Milano 1984), comunemente ricordato per «l'argomento della stanza cinese». Anche Searle è scandalizzato dalla facilità un po' incosciente con cui gli studiosi di Ia trattano l'intelligenza come fenomeno oggettivo, da smontare e ricostruire nei suoi procedimenti senza tener conto del suo intrinseco aspetto soggettivo, ovvero l'autoconsapevolezza del proprio procedere. Contro l'impostazione del test di Turing e contro l'idea che la sempre maggiore complessità e vastità dei programmi possa portare a veri e propri salti di livello, ad una diversa natura qualitativa del procedere per algoritmi, Searle propone il suo paradosso: anche se un sistema artificiale fosse così abile da superare il test di Turing e divenisse sotto tutti gli aspetti indistinguibile dal suo corrispondente biologico, tuttavia non sarebbe intelligente, perché non capirebbe quello che fa (pur se sostenesse di



capirlo e dimostrasse in tutti i modi, per esempio rispondendo disinvoltamente a qualsiasi domanda, di aver bene appreso la lezione). La parabola di Searle prevede che il programma in questione sia simulato da un essere umano chiuso in una stanza e dotato di tutte le regole necessarie per trattare dei simboli in ingresso, ad esempio dei caratteri cinesi, per elaborarli sintatticamente e per restituire in uscita una opportuna sequenza di simboli, in modo tale che, a fronte di domande in cinese fornite in input, l'uomo restituisca in output l'equivalente delle più naturali e spontanee risposte cinesi, pur senza sapere una parola di cinese. Per quanto l'impressione di un cinese dall'altra parte della parete sia quella di parlare con qualcuno che capisce la sua lingua, da nessuna parte, né nell'uomo, né nelle regole, né nella stanza nel suo complesso, c'è qualcosa che capisce il cinese. Similmente, un sistema che supera il test di Turing può non capire niente di ciò che dice, in cinese, in inglese o in qualsiasi altro linguaggio simbolico. Poiché, inoltre, in un programma di computer non c'è nulla di più di ciò che c'è nella stanza cinese, cioè soltanto regole e algoritmi per interpretarle (ed anzi c'è qualcosa di meno, perché l'hardware è ancora più stupido dell'omino che manipola i simboli), ecco che nemmeno alla fine dei tempi, quando un programma perfetto potrà generare l'illusione più completa, nemmeno allora esso potrà dirsi «intelligente».

È impossibile riassumere la varietà e la veemenza delle reazioni che l'apologo della «stanza cinese» ha suscitato nell'ambiente dell'intelligenza artificiale. L'ingenuità ostentata da Searle in fatto di programmi ha indispettito i ricercatori almeno quanto la loro ingenuità in fatto di «facoltà cosciente» aveva indispettito Searle, che alla riflessione su tale facoltà aveva dedicato tutte le sue fatiche filosofiche. L'operatività di una specie di impiegato postale e delle sue cataste di regole è lontana troppi ordini di grandezza anche da un banale sistema esperto; inoltre è discutibile che si possa realmente simulare la padronanza linguistica con la sola elabora-



zione sintattica dei simboli: anche Turing ipotizza che sia meglio «fornire alla macchina i migliori organi di senso che si possano costruire e poi insegnarle a capire e parlare una lingua umana». Qualunque sia l'idea che Searle ha dei programmi di computer, è difficile sostenere che una macchina dotata di organi di senso sia comunque e puramente «sintattica». I partigiani dell'Ia sembrano avere inoltre un asso nella manica: il cervello biologico non è forse una macchina? Se mettiamo un omuncolo a regolare e controllare tutte le sinapsi e le vicende dei miliardi di neuroni, non capirà ciò che avviene a livello di coscienza cerebrale più di quanto l'omino nella stanza capisca di cinese.

Il pensiero come «secrezione» biologica

Sorprendentemente, Searle è completamente d'accordo con l'ultima osservazione. Il cervello è effettivamente una macchina; un diavoleto che sorvegliasse tutti i neuroni non capirebbe nulla; ma attenzione, se in qualche modo il diavoleto tentasse una simulazione di tutto lo scambio sinattico, per esempio su un supporto fisicamente diverso, o su una diversa scala temporale, ecc., rischierebbe di far svanire l'autocoscienza del soggetto, pur riproducendone

per ipotesi tutte le reazioni fisiologiche e comportamentali. Per Searle, infatti, il pensiero cosciente è un prodotto biologico del cervello: per quel che ne possiamo sapere, soltanto la specifica composizione fisico-chimica della materia cerebrale è in grado di generare un pensiero soggettivo, delle sensazioni, l'effettiva impressione di piacere o di dolore. «Se sono in grado di capire l'inglese e ho altre forme di intenzionalità [...], per quanto ne sappiamo, è perché io sono un certo tipo di organismo con una certa struttura biologica (cioè chimica e fisica), e questa struttura, sotto certe condizioni, è causalmente capace di produrre percezione, azione, capacità di comprendere, di imparare, e altri fenomeni intenzionali». «Qualunque cosa sia l'intenzionalità, è un fenomeno biologico, e quindi è verosimile che sia causalmente dipendente dalla biochimica specifica delle sue origini, come la lattazione, la fotosintesi, o qualunque altro fenomeno biologico» (*Menti, cervelli e programmi*, pagg. 64 e 71). «Il cervello non rappresenta solo un'esemplificazione concreta di una configurazione formale o di un programma (è anche questo, certo), ma causa anche eventi mentali grazie a processi neurobiologici specifici. Il cervello è un organo biologico specifico e le sue proprietà biochimiche specifiche gli consentono di causare la coscienza e altri tipi di fenomeni mentali. [...] La

tesi dell'Ia forte è che un sistema arbitrario, sia esso fatto di lattine di birra, di chip di silicio o di carta igienica, purché attui il programma giusto, con gli ingressi giusti e le uscite giuste, non solo può avere pensieri e sentimenti, ma deve avere pensieri e sentimenti. Ebbene, questa è un'opinione fortemente antibiologica» («La mente è un programma?» in *Le Scienze* marzo 1990, pag.19). In breve, il cervello secerne pensiero come il fegato secerne bile: «i processi mentali sono parte della nostra storia biologica naturale tanto quanto la digestione, la crescita o la secrezione della bile» («La scienza cognitiva e la metafora del computer», in *Nuova Civiltà delle Macchine*, 1988 n. 1/2, pag. 61).

Il biologismo aprioristico di Searle è in genere respinto dai teorici dell'Ia. «Il fatto che intelligenza, capacità di comprensione, mente, consapevolezza, anima scaturiscano da una sorgente improbabile — un tessuto enormemente confuso di corpi cellulari, sinapsi e dendriti — è assurdo, eppure innegabile. Come questo possa creare un Io è difficile da capire, ma una volta che accettiamo quel fatto fondamentale, strano, disorientante, allora non dovrebbe sembrare più strano dare un Io a una tubatura d'acqua. [...] Searle dimentica di dire che una simulazione a tubi del cervello prenderebbe, diciamo, alcuni milioni di tubi con alcuni milioni di operai in piedi presso i rubinetti a girarli quando è necessario, e dimentica di dire che rispondere ad una domanda richiederebbe un anno o due» (Douglas Hofstadter, dibattito in *Menti, cervelli e programmi*, pagg. 114, 116). Ma il punto veramente decisivo è un'altro, e cioè che la tesi di Searle non ha niente a che vedere con l'argomentazione della stanza cinese, da cui sembrava scaturire la sua obiezione. È quindi possibile che, come spesso accade, Searle abbia ragione nella tesi di fondo, ma completamente torto negli argomenti usati per sostenerla. L'intuizione di Searle non riguarda specificamente i processi intelligenti e la polemica con l'Ia è del tutto occasionale e fuorviante. Egli sta pensando invece alla sensibilità biologica, alla capacità in

generale di provare piacere e dolore, e sostiene con una certa plausibilità che questa sia direttamente legata alle caratteristiche biologiche del sistema nervoso centrale. Biologismo a parte, si tratta del medesimo ragionamento svolto da Leibniz nella sua «Monadologia»: «Bisogna riconoscere — scriveva infatti Leibniz nel 1714 — che la percezione e quello che ne dipende è inesplicabile mediante ragioni meccaniche, cioè mediante figure e movimenti. Supposta una macchina la cui struttura faccia pensare, sentire e percepire; ed immaginato che s'ingrandisca, conservando le stesse proporzioni, in modo che vi si possa entrare come in un mulino; ciò fatto, nel visitarla internamente non si troverà altro che pezzi, i quali si spingono scambievolmente; e non mai alcuna cosa che possa spiegare una percezione. Cosicché questa bisogna trovarla nella sostanza semplice e non nel composto, o nella macchina». Searle non fa che attribuire la connaturata capacità sensibile della «monade» leibniziana alla cellula nervosa, senza per altro poter spiegare né come la cosa avvenga, né perché non accada con altri tipi di cellula o con composti più elementari. In ogni caso, la discussione verte sulla sensibilità soggettiva: non sulle attitudini logiche e razionali, come gli argomenti di Searle inducono a credere. Può darsi che la capacità di «sentire» qualcosa abbia a che vedere con la consapevolezza che supponiamo caratterizzi la mente umana nella sua più alta espressione, come «intelligenza di sé» e del proprio essere al mondo: ma certamente è qualcosa di diverso, che («per quel che ne sappiamo») condividiamo con tutto il regno animale, dal lombrico ai primati superiori. La capacità di provare «veramente» dolore o piacere, e non soltanto comportarsi «come se» si provasse, non ha bisogno di capire il cinese o di superare elaborati test linguistici, e non suppone alcuna forma di intelligenza raziocinante: abbiamo motivo di ritenere che appartenga anche alle prime fasi dell'evoluzione biologica e che anzi ne sia stato il motore selettivo più efficace. Ammesso che un lombrico provi dolore, pur con la sua rudimentale ca-

tena gangliare (la biologia delle sue cellule parrebbe averne tutta la necessaria «capacità causale»), e un pezzo di silicio o una libreria di regole invece no, non si vede che cosa possa aggiungere a questo proposito tutto l'argomento della stanza cinese e delle simulazioni linguistiche ed espressive. L'obiezione di Searle è già contenuta in questo semplice assunto: un essere biologico, per quanto elementare, è in grado di provare dolore; un hardware diverso, presumibilmente no.

Dalla ghiandola pineale di Cartesio in poi, l'ipotesi o la finzione di omuncoli nel cervello porta solo ad una «reductio ad infinitum» (come fa a sua volta la ghiandola o l'omuncolo a capire o a non capire?) perfettamente inutile: come può riguardare il cervello e lo stato mentale complessivo il fatto che qualcun altro, ispezionando o agendo sulle cellule cerebrali, capisca alcunché? Cosa cambia se per ipotesi l'uomo nella stanza capisse il cinese, oltre alle regole che continua fedelmente ad applicare? (Cosa cambia per il programma se anche il silicio, come nell'animismo delle favole, avesse «il giusto potere causale» per avere una coscienza?) Può darsi che un programma di computer digitale non riuscirà mai a superare un severo test di Turing: è una eventualità di tipo empirico, che ancora non siamo in grado di prevedere; ma se invece ci riuscisse, come Searle graziosamente concede per ipotesi alla sua stanza cinese, porrebbe comunque seri problemi di classificazione in ordine alla sua «intelligenza» e alla sua capacità raziocinante, e forse anche, sia pure in modo diverso che per il lombrico, in ordine alla sua «visione del mondo». Su questo punto, e sul lavoro dell'Ia in genere, la stanza cinese non incide in alcun modo, ed anzi finisce per veicolare una concezione di algoritmo ed un'idea della programmazione molto primitive e inadeguate, incapaci di cogliere i problemi più profondi che la materia impone.

L'argomento del lombrico

Val la pena invece esaminare il nocciolo più serio della posizione di Searle,

ciò la convinzione che un cervello biologico innanzitutto prova delle sensazioni, «sente» dolore e piacere. Che questa capacità, diciamo dal lombrico in su, sia strettamente fondata sulla natura biochimica dei processi biologici, è forse un'ipotesi indimostrabile, ma può essere ammessa come plausibile. Tuttavia, se accettiamo in qualche modo la millenaria distinzione tra senso e intelletto, tra passione e ragione, la impossibilità di riprodurre un «dolore artificiale» non comporterebbe una identica assurdità per il «ragionamento artificiale», o per la soluzione «ragionata» di problemi. È vero che le distinzioni non sono mai assolute, e che anche

biologico; con qualche caduta nei casi estremi, non è detto che non si possa conseguire la competenza linguistica per superare un buon test di Turing facendo a meno della «capacità del lombrico». Naturalmente, vi sono delle attività squisitamente intelligenti fondate sulla coscienza di sé, come la riflessione filosofica o la creazione poetica e artistica; esse però vengono in genere ben differenziate ed anzi contrapposte alla sensibilità animale, al biologismo del lombrico: danno più l'impressione dell'avvitarsi del pensiero su se stesso, di un epifenomeno che non ha precisi vantaggi biologici e che potrebbe essere un'eco, una risonanza del pensiero

te dell'Ia che si richiama al connessionismo.

Una sorpresa

Alcune delle osservazioni di Dreyfus e di Searle sono state fatte proprie dai ricercatori che lavorano sulle reti neurali; cosa pensano del connessionismo i due critici dell'Ia? A domanda diretta (si veda la «Intervista a J. Searle e a H. Dreyfus», *Office Automation*, gennaio 1988, pag. 51) Dreyfus ammette che «non vi sono motivi in linea di principio perché le connection machine non siano addestrate a comportarsi come se avessero buon senso». Searle invece precisa: «A me non interessano i comportamenti oggettivi, i comportamenti esteriori, ma la mente, i processi mentali più interni. E il punto è: la manipolazione di simboli non è sufficiente per produrre una mente o in una connection machine o in una macchina di von Neumann». Dreyfus cerca di convincere Searle che l'argomento della stanza cinese non vale più: infatti «la connection machine non usa simboli. Questo è il motivo per cui John ha bisogno di un nuovo argomento. [...] L'Ia tradizionale ha una struttura simbolica di cui le connection machine non hanno bisogno». Ma per Searle il fatto dell'hardware resta decisivo: se si costruiscono nuove macchine con diverse proprietà fisiche, forse si potrà anche ricreare il «potere causale» del cervello. «Però, in effetti, tutti i programmi esistenti di connection machine che conosco girano come i programmi di Ia delle macchine tradizionali di von Neumann. [...] Sono ancora programmi di Ia e di per se stessi non sono sufficienti ad avere comportamenti intenzionali». Successivamente, Searle ha ripreso le considerazioni di Dreyfus sulla innaturalità delle regole, e ora (1991) appare più cauto: «Una conseguenza inattesa della mia ricerca è che inavvertitamente sono arrivato a difendere, se questa è la parola giusta, il connessionismo. [...] Questo non è dire che i modelli connessionisti esistenti siano corretti — forse sono tutti sbagliati. Ma almeno non sono tutti ovviamente falsi o incoerenti come lo sono i modelli cogni-



le vicende dell'Ia suggeriscono un riavvicinamento tra l'intelligenza astratta e la capacità sensibile, l'informazione concreta, la creatività emotiva; se però uno spiraglio resta, esso si trova proprio all'estremo opposto della pura sensibilità del lombrico, in direzione del pensiero logico e dell'astrazione simbolica. Non stupisce più, allora, che di tutto il terreno del «mentale» si sia iniziato a ricostruire proprio l'aspetto dell'intelligenza, o che per esempio, come spesso si osserva, per il computer le cose più facili da fare siano quelle «più difficili», di più alto livello, mentre risulta più complicato emulare le cose che, se non proprio un lombrico, anche un topolino può compiere (per esempio, sfuggire all'agguato del gatto). Può darsi che vi sia una fetta molto rilevante di compiti «intelligenti» che non richiedano una autocoscienza in senso

astratto e simbolico più che un travestimento di una capacità sensitiva. La pretesa che l'Ia dia conto proprio di questi aspetti limite dell'intelligenza umana, a cui del resto nelle vicende alienanti e insensate del quotidiano è ben raro riuscire ad accedere, sembra un tantino esagerata. In questo senso una eventuale impossibilità «fisiologica» a ricostruire il senso di sé, fondata sull'argomento del lombrico, potrebbe non incidere molto sulla prospettiva della stessa Ia «forte»: una mente artificiale resterebbe forse «uno zombie senza intenzionalità» (Searle, «La coscienza e le scienze cognitive», *Sistemi Intelligenti*, agosto 1991, pag. 218), ma potrebbe benissimo risultare più intelligente di un buon numero di persone in carne ed ossa. L'argomento del lombrico, nonostante le apparenze, ha invece un effetto dirompente su quella corren-

tivisti tradizionali» («La coscienza e le scienze cognitive», pag. 216).

In realtà, l'apparente possibilismo di Searle dipende solo dalla sua indifferenza per i problemi concreti della ricerca e dal suo interesse esclusivo per le ragioni di principio. La sua tesi di fondo, che non ha niente a che vedere con la stanza cinese e col simbolismo ma che si richiama invece all'argomento del lombrico, pone di fatto al connessionismo una barriera praticamente invalicabile. Si tratta, se si vuole avere qualche speranza, di trovare un supporto fisico che, come l'analogo processo biochimico, per avventura abbia il «potere causale» di generare fin dall'inizio sensibilità soggettiva, senso del dolore, capacità di «sentire». Di fronte al programma neo-evoluzionista del connessionismo, che intende ricostruire dal basso le capacità mentali biologiche, l'istanza posta da Searle si presenta come passaggio obbligato.

Infatti, mentre si può sperare di trattare in qualche modo il pensiero astratto a prescindere dalla sensibilità animale, non si può immaginare di ripercorrere l'evoluzione biologica senza il motore di fondo costituito dalla ricerca del piacere e dalla fuga dal dolore. La sensibilità che ritroviamo anche negli organismi inferiori non è evidentemente il portato di schemi neurali straordinariamente complessi, ma potrebbe invece dipendere dalla struttura chimico-fisica del biologico, come sostiene Searle. Se questo è vero, non possiamo sperare di ottenerla alla fine del processo, complicando sempre più le maglie della rete: dovremmo riprodurla fin dall'inizio, nelle reti neuroniche più elementari. L'argomento del lombrico, se Searle ha ragione, è tanto più pressante quanto più lo sviluppo della mente artificiale è lasciata alla propria autoregolazione interna, alla equilibrizzazione «sub-simbolica» delle connessioni e delle comunicazioni.

La ricerca dell'equilibrio non ha in sé una spinta propulsiva se non è integrata dalla imposizione esplicita di finalità e di obiettivi significanti, oppure, necessariamente, da una spontanea tensione a rifuggire le situazioni spiacevoli e rincorrere quelle piacevoli. Nella vita

biologica elementare, il «principio del piacere» è l'unica molla e l'unica condizione che permette ad un sistema «autopoietico» di sopravvivere e di evolvere.

Una rete connessionista può risolvere problemi impostati dall'esterno, che hanno cioè una valenza simbolica almeno nei nodi di input e di output; perciò può funzionare egregiamente come supporto a «un sistema fisico di simboli», per esempio a un sistema esperto che la utilizzi per il riconoscimento che la utilizzi per il riconoscimento di un sistema senza che le sia esplicitamente fornito l'insieme dei valori e dei fini da conseguire. L'unica alternativa nota richiede appunto la capacità di percepire dolore e piacere, come motivi di fuga e di ricerca: per questo l'argomento del lombrico risulta qui molto più drammatico, anche sotto l'aspetto immediatamente pratico.

È d'altra parte esperienza ormai ricorrente, nella sperimentazione connessionistica, raggiungere fasi di apprendimento in cui la rete «collassa», per esempio non rispondendo più agli stimoli, o viceversa rispondendo in modo indifferenziato e incontrollabile. Se le finalità non sono pressantemente imposte da esigenze esterne, la rete «non ha motivo», per così dire, di fornire prestazioni utili.

Mentre possiamo immaginare uno «zombie» intelligente nel senso dell'Ia classica, uno «zombie» interamente connessionistico che non soddisfi all'argomento del lombrico, cioè che non senta piacere e dolore, sarà paragonabile al bambino autistico, potenzialmente intelligente, ma che non ha motivo di agire e pensare, che è indifferente a ciò che accade e che indifferente si procura danni e lesioni. Il programma evolutivo degli ultimi connessionistici sembra così essere messo fortemente in crisi, quanto più per l'appunto essi aspirano a emulare gli autonomi meccanismi dell'evoluzione biologica; è perlomeno curioso che siano proprio loro i più fedeli sostenitori della stanza cinese, delle critiche alla Searle e, indirettamente, dell'argomento del lombrico.

Qual è la differenza tra la fragola e l'elefante?

Più che dallo studio degli organismi inferiori e dei fondamenti per una «vita artificiale», le critiche più costruttive all'intelligenza artificiale vengono dall'esame differenziato delle funzioni cerebrali superiori, per esempio come si manifestano nelle malattie mentali o nelle alterazioni del comportamento normale. A volte la selettività della lesione genera una specie di comportamento astratto, isolando un aspetto funzionale e mostrandolo nel suo (innaturale ma istruttivo) isolamento. Gli esempi più noti al pubblico sono dovuti alla abilità letteraria di Oliver Sacks, autore di *Risvegli* e de *L'uomo che scambiò sua moglie per un cappello* (Adelphi 1987 e 1986). Uno dei casi più significativi è proprio quello che dà il titolo al secondo volume: una persona colta e intelligente è affetta da un processo degenerativo nelle parti visive del cervello che disturba il normale riconoscimento delle immagini. Il signor P., pur «ricevendo» tutta l'informazione visiva e tutti i particolari dell'immagine, non è in grado di averne una visione d'insieme, di darne una interpretazione complessiva. Non riconosce infatti i volti e le cose più familiari, ma è costretto a dedurne faticosamente l'identità ricostruendo i dettagli e le connessioni più specifiche. «Egli non riusciva a formulare un giudizio cognitivo, pur essendo in grado di produrre numerose ipotesi cognitive. Un giudizio è intuitivo, personale, comprensivo e concreto: noi vediamo come stanno le cose, in relazione tra di loro e con noi stessi. Era proprio questa capacità di vedere, di stabilire relazioni che mancava al dottor P.» (*L'uomo che scambiò ecc.*, pag. 38). Sacks gli sottopone fotografie di volti a lui ben noti: «Si può dire che non riconobbe nessuno: né i familiari, né i colleghi, né gli allievi, né se stesso. Riconobbe un ritratto di Einstein perché ne individuò i caratteristici capelli e baffi; lo stesso avvenne con un paio di altre persone. Ah, «Paul!» disse quando gli mostrai un ritratto del fratello «Quella mandibola squadrata, quei grossi denti! Lo riconoscerai dovunque!». Ma era Paul che ri-

conosceva, o erano alcuni tratti che gli permettevano di azzardare un nome con buona probabilità di aver ragione? [...] Egli si accostava a questi volti, anche quelli delle persone a lui più vicine e care, come a un rompicapo o a test astratti. Non stabiliva nessun rapporto con loro, non li vedeva. Nessun viso gli riusciva familiare, non lo vedeva come un tu, ma lo identificava come un semplice insieme di tratti, un esso. Vi era quindi gnosis formale, ma nessun accenno a una gnosi personale» (pag.

pura elaborazione concettuale che non può appoggiarsi a sicure intuizioni d'insieme ed è costretta a ricostruire le situazioni a partire dai particolari che le sono dati. «Certo, il cervello è una macchina e un elaboratore. Ma i processi mentali, che costituiscono il nostro essere e la nostra vita, non sono soltanto astratti e meccanici, sono anche personali; e in quanto tali implicano non solo la classificazione e l'ordinamento in categorie, ma anche una continua attività di giudizio e di senti-

Il pensiero astratto

Ciò nonostante, il caso del signor P. indica anche qualche cos'altro. Suggerisce che la nostra mente, accanto al giudizio e all'intuizione d'insieme, usa effettivamente l'intelligenza astratta, che non è dunque una chimera e un'indebita astrazione. Il ragionamento deduttivo ha un suo peso reale, presumibilmente là dove l'intuizione può essere ingannata e il giudizio basato sull'ovvio non è più sufficiente; esso inoltre può rimanere inalterato, anche quando la capacità intuitiva viene a mancare. Questa relativa indipendenza dei due aspetti, mentre costituisce un monito per chi vuole ricostruire l'interezza del mentale, suggerisce anche la possibilità di operare secondo linee distinte: il progetto di una «intelligenza artificiale» in senso classico rimane plausibile,



31). P. è costretto a supplire con l'intelligenza e la deduzione alla sua carenza di intuizione: «E questo che cos'è?» chiesi sollevando un guanto. «Posso esaminarlo?» disse. Me lo prese di mano e si mise a studiarlo come aveva fatto con le forme geometriche. «Una superficie continua» annunciò infine «avvolta su se stessa. Dotata...» esitò «di cinque estensioni cave, se così si può dire». «Sì» dissi cautamente «Lei mi ha fatto una descrizione. Ora mi dica che cos'è». «Un qualche contenitore?» «Sì» dissi «e che cosa potrebbe contenere?» [...] «Ci sono varie possibilità. Potrebbe essere un portamonete, per esempio, per monete di cinque valori diversi. Oppure...» [...] In seguito se lo infilò per caso: «Dio mio!» esclamò. «È un guanto!» (pag. 32-33). Il signor P. incarna in qualche modo, almeno nei confronti della interpretazione delle immagini, l'intelligenza astratta, la

mento. Se ciò va perduto, finiamo, come il dottor P., per assomigliare a degli elaboratori» (pag. 40). Rispetto alla pienezza della coscienza, i tentativi maldestri del signor P. rappresentano una patetica caricatura, che offre lo spunto a Sacks per una critica al cognitivismo psicologico: «Per una sorta di comica e spaventosa analogia, la neurologia e la psicologia cognitive odierne presentano una forte somiglianza proprio col povero dottor P.! Come lui abbiamo bisogno del concreto e del reale; e come lui non ce ne accorgiamo. Le nostre scienze cognitive soffrono anch'esse di una agnosia essenzialmente simile a quella del dottor P. Il quale può dunque servire da monito e da parabola, mostrandoci che cosa succede a una scienza che rifugge dal giudizio, dal particolare, dal personale e diventi interamente astratta e computazionale» (pag. 40).

BIBLIOGRAFIA

Hubert Dreyfus: *Che cosa non possono fare i computer*, Armando, 1988.

John Searle: *Menti, cervelli e programmi* (con osservazioni di AA.VV.), Clup-Clued, 1984.

John Searle: «La mente è un programma?», *Le Scienze*, marzo 1990, pagg. 16-21.

John Searle: «Le insoddisfazioni del materialismo» e *La coscienza, l'inversione della spiegazione e le scienze cognitive*, in *Sistemi Intelligenti* agosto 1991, pagg. 171-220.

Oliver Sacks: *L'uomo che scambiò sua moglie per un cappello*, Adelphi, 1986.

Oliver Sacks: *Risvegli*, Adelphi, 1987.

Douglas Hofstadter e Daniel Dennet: *L'io della mente*, Adelphi, 1987.

Terry Winograd e Fernando Flores: *Calcolatori e conoscenza*, Mondadori, 1987.

Joseph Weizenbaum: *Il potere del computer e la ragione umana*, Edizioni Gruppo Abele, 1987.

Luciano Bazzocchi: «La mente dell'uomo e l'intelligenza del computer», *Office Automation*, gennaio 1988.

Sul numero 1/2 1988 di *Nuova Civiltà delle Macchine*, dedicato alla discussione su «Conoscenza e intelligenza artificiale», si vedano in particolare: Alberto Oliverio: *Intelligenze biologiche e intelligenze artificiali*, pagg. 36-42;

John Searle: *La scienza cognitiva e la metafora del computer*, pagg. 53-61;

Hubert Dreyfus: *Si può accusare Socrate di cognitivismo?*, pagg. 62-72;

Luciano Bazzocchi: *Intelligenza artificiale e sistemi esperti*, pagg. 113-120.

Heinz Pagels (a cura di): *La cultura dei computer*, Boringhieri, 1989.

Paul e Patricia Churchland: «Può una macchina pensare?», *Le Scienze*, marzo 1990, pagg. 22-27.

sia pure da integrare, possibilmente, con altre funzionalità di tipo diverso. L'ipotesi proposta da Marvin Minsky nella sua Società della mente non è dopo tutto così irrealistica e astratta: «Il cervello umano è una vasta società organizzata, composta di molte parti diverse. Dentro il cranio dell'uomo sono stipati centinaia di tipi diversi di motori e organizzazioni, meravigliosi sistemi evolutisi e accumulatisi nel corso di centinaia di milioni di anni. [...] Queste parti di mente devono convivere con le

altre, in un rapporto che è a volte di collaborazione, ma più spesso di conflitto» (Adelphi 1989, pag. 20).

Si può ammettere, con Sacks, che la componente più propriamente intelligente non sia la più importante, dal punto di vista dell'evoluzione: «Un animale, o un uomo, può fare benissimo a meno dell'«atteggiamento astratto», ma perirà rapidamente se privato del giudizio» (op. cit. pag. 40). Tuttavia, il signor P. prova che l'atteggiamento astratto ha una sua realtà autonoma,

anche là dove manca la capacità di giudizio. Il riduzionismo, cioè l'ipotesi che la mente si possa comprendere distinguendola in parti e in funzioni relativamente autosussistenti, è ancora un utile strumento esplicativo; il progetto di una «intelligenza artificiale», necessariamente parziale rispetto ad una «mente artificiale» pienamente adeguata, resta una prospettiva concretamente perseguibile. ■

Luciano Bazzocchi

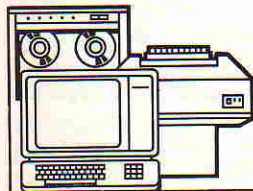
(4 - continua)

TELECOMUNICAZIONI E TRASMISSIONE DATI

PROGETTAZIONE SISTEMI
PROGETTAZIONE RETI
CONSULENZA
ASSISTENZA SISTEMISTICA
FORNITURE PERSONALIZZATE
INSTALLAZIONI «CHIAVI IN MANO»
MANUTENZIONE
RACK
APPARATI T.D.
BORCHIA MULTIPLA
DIRAMATORI
COMMUTATORI
ACCESSORI
CAVI



FILTE sud S.R.L.
TELECOMUNICAZIONI
E TRASMISSIONE DATI



Via Villafranca, 15
00040 CECCHINA (Roma)
Tel. (06) 93.43.251
Telefax (06) 93.42.016
Telex 610561 FILTES I



Dossier del mese
Hand held computer
Inchiesta
Ultime novità dalle Lan (2ª parte)

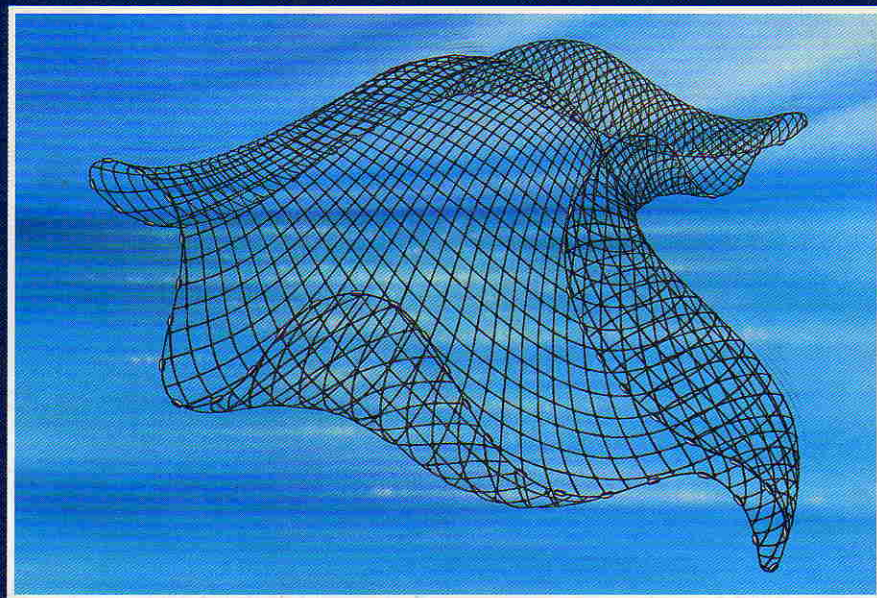
l'organizzazione

l'automazione e le comunicazioni dell'ufficio

office[®] automation

SOIEL INTERNATIONAL - 20125 MILANO, VIA MARTIRI OSCURI 3 - MENSILE - ANNO DODICESIMO - N. 1, GENNAIO 1992 - SPED. IN ABB. POST. GR. III/70

ABBIAMO GETTATO LE RETI NEL MARE DELLA COMUNICAZIONE



ERICSSON SIELTE
reti per comunicare

ERICSSON 