

FILOSOFIA, INTELLIGENZA ARTIFICIALE E APPRENDIMENTO

Alberto Giovanni Biuso

1. Che cos'è l'Intelligenza Artificiale?

L'Intelligenza Artificiale come ambito autonomo di ricerca nasce con la conferenza di Dartmouth del 1956. Fin dall'inizio si delinearono due ipotesi assai diverse fra di loro. Il programma dell'IA *debole* guardò a questa prospettiva in modo sostanzialmente operativo e strumentale, come a una maniera particolarmente efficiente di dotare l'umanità di macchine in grado di eseguire con efficacia lavori non solo di routine e molto pesanti – compito che le macchine hanno svolto egregiamente almeno a partire dalla Rivoluzione Industriale – ma anche di *simulare* attività intelligenti in quei settori nei quali non basta la semplice forza meccanica. L'IA *forte* rappresentò un programma assai diverso rispetto al primo, ritenendo che, anche se molto difficile, non sia per principio impossibile la creazione di menti artificiali, dotate di una qualche forma di *coscienza separata* dalla struttura biologica del corpo umano.

Gli obiettivi principali dell'attuale ricerca sull'*Artificial Intelligence* possono essere così sintetizzati:

- 1) tentare di portare i computer nella classe delle entità intelligenti;
- 2) simulare su un computer il processo conoscitivo, il comportamento, le reazioni cerebrali umane, allo scopo di comprendere meglio la nostra mente;
- 3) studiare la natura, l'identità e la differenza fra l'intelligenza umana e quella di altre entità sia reali – macchine, gli altri animali – che ipotetiche: alieni.

Tali obiettivi si intrecciano con alcune finalità di carattere più generale, come ad esempio:

1) un orientamento *teorico*, indirizzato soprattutto alla comprensione e spiegazione di che cosa sia la mente;

2) un indirizzo *tecnologico*, volto a costruire computer che possano sostituire o integrare gli umani in una svariata serie di compiti;

3) uno scopo *commerciale*, interessato a produrre, distribuire, vendere merce di largo consumo.

Si potrebbe dire che l'*Artificial Intelligence* occupi uno spazio intermedio e fortemente variabile fra la domanda filosofica, psicologica, neurologica sull'identità dell'*Intelligenza* e le ricerche tecnologiche, informatiche e commerciali sulla *Robotica*. Sono quindi numerosi e importanti i settori in essa coinvolti e un posto particolare è occupato dalla filosofia della mente. Poniamo quindi la domanda di fondo nella quale filosofia e Intelligenza Artificiale si intersecano: che cosa significa pensare? In che cosa consiste l'identità del soggetto umano rispetto agli altri animali e in relazione alle macchine che elaborano dati? È possibile un'evoluzione della nostra specie verso l'Intelligenza Artificiale? E, di converso, sino a che punto i calcolatori possono e potranno assumere modalità umane di pensiero? Dove, in quale dimensione, collocare il possibile punto di incontro tra l'artificializzazione dell'umano e l'umanizzazione delle macchine? Sembra anzitutto che per avvicinarsi *davvero* al pensiero un cervello artificiale non dovrebbe limitarsi a *eseguire* programmi ma dovrebbe *produrre* eventi mentali tramite processi di tipo neurobiologico. A fondamento del problema rimane dunque la domanda sui rapporti tra mente e cervello. Può l'intelligenza essere separata dal cervello? Le correnti materialistiche lo negano decisamente. Jean-Pierre Changeux – in polemica implicita con posizioni come quelle di Hartmann e di Lorenz – propone infatti un'epistemologia marcatamente riduzionistica, che abbia a fondamento «il metodo cosiddetto di riduzione, o meglio ancora di ricostruzione, che tutti usiamo nelle scienze sperimentali [e che] consiste nella ricerca di una spiegazione a un livello inferiore a quello che si vuole spiegare»¹.

¹ J.P. Changeux e A. Connes, *Pensiero e materia (Matière à pensée, 1989)*, trad. di C. Milanese, Torino, Bollati Boringhieri, 1991, p. 144.

Al riduzionismo fisicalistico si oppongono coloro che distinguono tra l'attività del pensare e il supporto mediante il quale essa viene esercitata. Descartes rifiuta persino l'idea che la vita sia un requisito fondamentale della coscienza. Per lui, la *res cogitans* è una forma d'essere ontologicamente diversa dalla fisicità e quindi integralmente separabile dall'estensione corporea. Gli organismi in quanto tali possiedono lo stesso statuto delle macchine. Se è quindi possibile un corpo privo di autocoscienza, sarà anche plausibile una coscienza viva e operante su un sostrato diverso rispetto all'organismo. Le macchine potrebbero quindi provare sensazioni diverse dalle nostre, ma non per questo sarebbero impossibilitate a possedere una qualche forma di sensibilità.

La tesi funzionalistica – uno dei cui maggiori esponenti è stato per molto tempo Putnam – trasferisce i termini nel linguaggio della programmazione, assimilando il cervello all'*hardware* e la mente al *software*, e rendendo in questo modo la mente autonoma dalla base materiale costituita dalla componente fisica e chimica del cervello. Niente, infatti, impedirebbe di trasferire un programma da un supporto a un altro². Seppur con modalità e obiettivi assai diversi, anche Popper distingue il «mondo 1» degli oggetti materiali (alberi, rocce, sedie, i corpi), il «mondo 2» delle sensazioni, dei pensieri e degli stati emotivi del singolo soggetto umano, il «mondo 3» delle teorie, delle narrazioni, delle produzioni artistiche, delle argomentazioni logiche, delle spiegazioni scientifiche. Fra questi tre mondi l'interazione è profonda e costante *in ogni direzione*: il mondo 2 è radicato nel corpo (mondo 1) ma è anche costantemente prodotto dal mondo 3 della cultura mentre contribuisce alla sua elaborazione. Le teorie filosofiche, le ipotesi scientifiche, le religioni, le forme d'arte incidono fortemente sugli oggetti materiali at-

² «Il funzionalismo stabilisce un'analogia tra la relazione che intercorre tra gli stati mentali e i processi del cervello, da una parte, e la relazione tra gli stati logici e le operazioni elettrofisiche della macchina dall'altro» (G. Lolli, «La filosofia della mente», in P. Rossi e C.A. Viano, *Storia della filosofia*, vol. VI, tomo 2, Roma-Bari, Laterza, 1999, p. 741). Anche Hofstadter pone a fondamento della propria riflessione la stretta analogia fra mente/software e cervello/hardware e ritiene che da un sempre più veloce sviluppo quantitativo (memoria e velocità) e qualitativo dei computer possa emergere qualcosa di simile alla coscienza (oltre a *L'io della mente*, si veda Gödel, Escher, Bach: *an Eternal Golden Braid*, 1979; trad. it. di G. Trautteur, Milano, Adelphi, 1990, in particolare il cap. XI, «Cervelli e pensieri», pp. 365-396, e il XX, «Strani Anelli e Gerarchie Aggrovigliate: il cuore dell'IA», pp. 741-780).

traverso la mediazione delle singole menti, delle loro capacità emotive, riflessive, problematiche, razionali, operative. E quindi «quello che noi chiamiamo “mondo 2” – il mondo della mente – diviene sempre più, al livello umano, il legame fra il mondo 1 e il mondo 3»³. Importante è anche l'ipotesi dell'innatismo cerebrale e cognitivo dal quale la mente umana sarebbe caratterizzata: linguisti e antropologi come Chomsky e Lévi-Strauss ritengono che la mente umana strutturi l'esperienza a partire da modelli innati.

Una conclusione – per quanto provvisoria – potrebbe essere la seguente: anche la filosofia della mente non ha una direzione univoca di ricerca proprio perché il suo oggetto si è differenziato in una serie di ambiti e di indagini tutte in qualche modo al confine fra il biologico e il computazionale. L'apertura metodologica, la pluralità degli approcci, la criticità verso ogni risultato acquisito sono alcuni degli elementi perenni della filosofia. Essi caratterizzano anche la ricerca filosofica sulla mente e sull'Intelligenza Artificiale.

2. Intelligenza umana e Intelligenza Artificiale

Lo stretto rapporto fra il mentale, il biologico e l'artificiale induce quindi a chiederci che cosa sia l'intelligenza naturale. È questa una domanda che può avere numerose e diverse risposte. *Intelligenza* vuol dire comprensione dell'ambiente in cui si vive e il conseguente migliore adattamento possibile a esso; significa capacità di intuire e di sapere affrontare la complessità delle situazioni esistenziali; comporta l'attitudine a riflettere a fondo prima di prendere delle decisioni che implicino conseguenze a volte anche irreversibili; coincide con un atteggiamento

³ K.R. Popper, *La conoscenza e il problema corpo-mente (Knowledge and the Body-Mind Problem. In defence of interaction*, 1994), trad. it. di F. Laudisa, a cura di M.A. Notturmo, Bologna, Il Mulino, 2002, p. 187. E ancora: «Non possono esservi processi di pensiero senza contenuti di pensiero, e i contenuti di pensiero appartengono al mondo 3» (*ibid.*, p. 153); «la sequenza dei numeri naturali 1, 2, 3... è un'invenzione umana, un'invenzione linguistica. La distinzione fra numeri pari e numeri dispari, invece, non è un'invenzione, ma una scoperta che facciamo all'interno di un oggetto del mondo 3 – la sequenza dei numeri naturali – che noi stessi avevamo prodotto e inventato» (K.R. Popper, *Tre saggi sulla mente umana*, a cura di A. Benini, Roma, Armando, 1996, p. 51).

mento equilibrato rispetto alle scelte etiche di fronte alle quali la vita continuamente ci pone; indica la razionalità in atto nel calcolo matematico e nella comprensione oggettiva del mondo; è comprensione non solo degli oggetti ma anche dello strumento che su di essi indaga; è capacità di trovare soluzioni diverse a problemi ricorrenti; è facoltà di apprendimento rispetto al vissuto e di infinita autocorrezione.

Che cosa si intende, di contro, per *Artificiale*? Quest'ultimo si colloca su un piano ontologico peculiare, diverso sia dal *Naturale* che dal *Tecnologico*. La Natura produce degli enti senza l'apporto dell'attività umana. La Tecnologia rappresenta il tentativo di conoscere, controllare, dominare e utilizzare gli enti e le forze naturali a vantaggio della nostra specie con l'ausilio di macchine e apparati inventati dall'uomo. L'Artificiale presuppone un esemplare esistente in natura che faccia da riferimento e che si tenti in qualche modo di riprodurre. Non ha senso, ad esempio, parlare di un «televisore artificiale» in quanto l'oggetto televisore è un prodotto tecnologico che in natura non esiste. Ugualmente distante dagli enti naturali come dagli artefatti, dal biologico come dall'automazione, l'Artificiale sembra, dunque, «una terza realtà fatta di tecnologia ispirata alla natura, una realtà, se si vuole, *naturoide*»⁴. La tecnologia convenzionale cerca di *dominare* la natura mediante delle invenzioni; il suo eroe potrebbe essere Prometeo che ruba e utilizza l'energia naturale allo scopo di affrancare la specie umana dagli dèi e renderla così padrona di se stessa e signora della natura. La tecnologia dell'artificiale tenta – invece – di *riprodurre* a vari livelli gli enti minerali, vegetali, animali, attraverso delle strategie imitative; un suo simbolo è quindi Icaro, che imita le creature alle quali la natura ha offerto il dono del volare.

Il rapporto tra λόγος e φύσις, tra la mente e la natura, può essere rivisto alla luce degli sviluppi continui delle macchine per pensare, con le loro capacità di calcolo sempre più imponenti, con velocità di elaborazione impensabili solo fino a qualche anno fa, con successi tali da far apparire prossimo il momento nel quale le macchine supereranno il test di Turing, saranno in grado cioè di replicare alle domande poste da

⁴ M. Negrotti, *Artificiale. La riproduzione della natura e le sue leggi*, Roma-Bari, Laterza, 2000, p. 12.

un interlocutore in un modo che renda le loro risposte indistinguibili da quelle fornite da un essere umano. Com'è evidente, però, si tratta di un test pensato in forma totalmente antropomorfa e addirittura comportamentistica: intelligente sarà quella macchina che fornirà delle risposte-reazioni analoghe alle risposte-reazioni di un essere umano posto nella medesima situazione⁵. Dalla stessa congettura nasce la metafora – sostenuta dai difensori e negata dai detrattori dell'IA – secondo la quale la mente sta al cervello come il software sta all'hardware. Da questi presupposti dipendono anche due grandi paradigmi come il dualismo e il materialismo. Per il primo la mente ha uno statuto ontologico diverso dal corpo. La metafisica cartesiana è uno degli esempi più chiari e più famosi di tale paradigma. Per essa *res cogitans* e *res extensa* sono due sostanze diverse la cui stretta corrispondenza comportamentale rimane però un problema che ha generato i più vari tentativi di spiegazione. Il materialismo risolve il dilemma eliminando uno dei suoi due corni tramite la tesi che il mentale altro non sia che una forma del fisico, che il pensiero consista nelle sinapsi attraverso le quali i neuroni si scambiano fra di loro l'energia che li fa vivere. Come accennato, il paradigma funzionalista ha cercato di superare le unilateralità del dualismo e del materialismo riconoscendo «la possibilità che sistemi così diversi fra loro come esseri umani, macchine calcolatrici, spiriti disincarnati possano tutti avere stati mentali. Nella concezione funzionalistica la psicologia di un sistema non dipende da ciò di cui è fatto (cellule viventi, metallo o energia spirituale), ma dal modo in cui è assemblato ciò che lo compone»⁶.

Neppure quest'ultimo paradigma spiega però il paradosso che ha progressivamente privato di forza il programma di ricerca dell'IA. Un paradosso che si può esprimere in questo modo: le macchine si vanno dimostrando capaci di affrontare e risolvere i problemi che coinvolgono le capacità intellettive più elevate, il ragionamento astratto e mate-

⁵ Il test venne proposto da Alan Turing in un famoso articolo del 1950 dal titolo *Computing machinery and intelligence*, in «Mind» 59 (236), pp. 433-460, trad. it. di G. Longo, *Calcolatori e intelligenza in L'io della mente* cit., pp. 61-74. La descrizione del contesto e una considerazione critica del test di Turing si trovano in R. Penrose, *La mente nuova dell'imperatore (The Emperor's new Mind)*, 1989), trad. it. di L. Sosio, Milano, Rizzoli, 2000, pp. 24-31.

⁶ J.A. Fodor, *Il problema mente-corpo*, in «Mente e macchina. Quaderni Le Scienze» 66, a cura di G. Lolli, giugno 1992, p. 19.

matico, ma si rivelano ostinatamente inadeguate a decifrare e persino a percepire i problemi più banali della vita quotidiana e del cosiddetto buon senso. Detto in altro modo: un calcolatore esegue perfettamente le istruzioni contenute in un algoritmo, ma riconosce con estrema difficoltà le forme nello spazio. Quest'ultimo problema, infatti, è di tipo non computazionale, ma aleatorio. Riuscire, ad esempio, a riconoscere un volto umano fra tanti comporta la necessità di memorizzare *tutte* le possibili alternative in modo da trarre – dal database della memoria – una soluzione che valga solo per quel caso specifico e non per tutti gli altri volti, che possono essere fra di loro somiglianti, ma che rimangono naturalmente diversi.

Molti fra i sostenitori dell'IA pensano di superare questi ostacoli – indubbiamente assai ardui – attraverso la sostituzione del calcolo seriale con quello *parallelo*; la potenzialità delle macchine starebbe in un modo altro di *pensare*, non sequenziale ma reticolare: «oggi, dopo decenni di tentativi infruttuosi di insegnare ai circuiti elettronici come pensare in modi cartesiani sequenziali sull'architettura sequenziale cartesiana, questa comunità [dei ricercatori sull'IA] si è mossa nel senso di adottare come strumenti standard le intermatematiche parallele»⁷. Il presupposto teorico è che la principale differenza fra i calcolatori e il cervello consista nel fatto che quest'ultimo non è una macchina seriale digitale che esegue un calcolo alla volta, ma consiste in miliardi di calcolatori di potenza limitata e ciascuno inconsapevole di sé – i neuroni – collegati però in parallelo e cioè in modo tale che il loro lavoro complessivo produca comprensione, soluzione, progresso conoscitivo. Un'altra ipotesi si affida ancora alla pura forza del calcolo e ritiene che dall'accrescimento quantitativo possa prima o poi *emergere* qualcosa di simile a un'autocoscienza. Un diverso approccio, per quanto somigli a quello funzionalista, risulta più complesso e modifica la metafora computazionale della mente/hardware nel senso delle *reti neurali* che imitino quella naturale costituita dal cervello. Una rete neurale è sostanzialmente una «simulazione al computer di gruppi di cellule nervose, che possono essere addestrate o program-

⁷ J. Bailey, *Il postpensiero. La sfida dei computer all'intelligenza umana* (After Thought, 1996), trad. it. di L. Sosio, Milano, Garzanti, 1998, p. 256.

mate per apprendere schemi di comportamento»⁸. Le reti neurali sembrano avere una migliore capacità di generare soluzioni attraverso l'uso controllato della casualità e degli algoritmi genetici, vale a dire algoritmi che producono diverse soluzioni per poi selezionare quelle che offrono i risultati migliori.

Le soluzioni indicate – il calcolo parallelo, l'emersione della coscienza dalla forza computazionale, le reti neurali, gli algoritmi genetici – sembrano collocarsi in una posizione intermedia fra il permanere del modello antropomorfo e la ricerca di un paradigma effettivamente diverso e adeguato alla novità sostanziale che l'IA rappresenta nell'ambito della gnoseologia, dei sistemi cognitivi e della stessa ontologia. Affinché dei robot diventino intelligenti sarà necessario, infatti, riprodurre in loro uno sviluppo non solo ontogenetico ma anche filogenetico, poiché anche le creature di silicio hanno bisogno di un corpo che le immerga nell'ambiente dal quale trarre le più preziose informazioni, quelle – apparentemente semplici – che consentono agli umani di muoversi nel mondo con estrema facilità, prontezza, inventiva.

L'interazione fra il silicio e i tessuti organici sembra anche un tentativo di superare una delle più forti obiezioni epistemologiche che l'IA abbia ricevuto: la «stanza cinese» di Searle. Il filosofo americano rifiuta in maniera decisa la metafora computazionale e tramite il suo esperimento mentale enuncia alcuni assiomi e delle conclusioni che hanno alla base la distinzione posta da Morris fra la sintattica, la semantica e la pragmatica. L'esperimento consiste nell'immaginare che una persona totalmente ignara della lingua cinese venga chiusa in una stanza nella quale trova soltanto un manuale in inglese che illustri la corrispondenza fra gli ideogrammi cinesi e il lessico e la sintassi dell'inglese. Attraverso tale manuale, il soggetto sarà in grado di rispondere alle domande in cinese che gli vengono trasmesse da una apertura della stanza e saprà anche trovare le risposte corrette. Nessuno, tuttavia, potrà sostenere che questa persona conosca davvero il cinese. Per Searle, un computer che fosse anche in grado di superare il test di Turing somi-

⁸ J. Brown, *Menti, macchine e multiverso. Alla ricerca del computer quantistico (Minds, Machines, and the Multiverse. The Quest for the Quantum Computer, 2000)*, trad. it. di S. Frediani, Torino, Einaudi, 2003, p. 142.

glierebbe in realtà al soggetto dell'esperimento in questione. In sostanza, secondo Searle la computazione è una sorta di produzione di segni per mezzo di segni, senza che tali connessioni aggiungano alcuna informazione sul contenuto semantico dei segni utilizzati e sulla relazione tra i segni, l'interprete e il contesto generale che dà significato alle operazioni sintattiche. Poiché si tratta di un'obiezione rivolta alla natura stessa della computazione, essa varrebbe anche nel caso in cui il calcolo seriale fosse sostituito da quello parallelo. La riduzione del pensare alle sue componenti sintattiche si fonda inoltre, a parere di Searle, su alcuni presupposti fortemente antibiologici che riprendono in pieno il paradigma dualista della relazione mente/corpo, paradigma incapace di comprendere «che la mente è un fenomeno biologico al pari della digestione. La mente, si suppone, è qualcosa di astratto e formale, non fa parte di quella "roba umida e appiccicaticcia" contenuta nella nostra testa»⁹. Il dualismo ontologico coniugato alla gnoseologia comportamentistica del test di Turing genererebbe l'illusione di poter creare delle menti attraverso la compilazione di programmi, proprio perché la mente come *software* sarebbe del tutto indipendente dal tipo di *hardware* sul quale il programma viene fatto girare. E, in effetti, condizione del programma dell'IA forte è l'indipendenza della mente da un particolare tipo di supporto biochimico. È opportuno, a questo punto, riassumere gli assiomi e le conclusioni che Searle trae dall'esperimento mentale della stanza cinese:

Assioma 1. *I programmi di calcolatore sono formali (sintattici).*

Assioma 2. *La mente umana ha contenuti mentali (una semantica).*

Assioma 3. *La sintassi di per sé non è condizione essenziale, né sufficiente, per la determinazione della semantica.*

Conclusione 1. *I programmi non sono condizione essenziale né sufficiente perché sia data una mente.*

Assioma 4. *Il cervello causa la mente.*

⁹ J.R. Searle, *La mente è un programma?*, in «Mente e macchina» cit., p. 10; si veda inoltre *Menti, cervelli e programmi*, trad. it. in *L'io della mente* cit., pp. 341-360. Anche Hofstadter sembra ammettere «che le macchine non avranno mai un'intelligenza di tipo umano finché non avranno qualcosa di simile a quell'umidità e scivolosità biologica. Non intendo umido in senso letterale; la scivolosità potrebbe essere nel software» (in *L'io della mente* cit., p. 94).

Conclusione 2. *Qualunque altro sistema in grado di causare una mente dovrebbe possedere poteri causali (almeno) equivalenti a quelli del cervello.*

Conclusione 3. *Qualunque sistema artificiale in grado di produrre fenomeni mentali; ossia qualunque cervello artificiale, dovrebbe essere in grado di riprodurre gli stessi poteri causali specifici del cervello umano e non potrebbe farlo soltanto svolgendo un programma formale.*

Conclusione 4. *Il modo in cui il cervello umano produce effettivamente i fenomeni mentali non può ridursi solamente allo svolgimento di un programma al calcolatore*¹⁰.

Fino a che punto lo stesso Searle estende l'applicabilità di tali assiomi e conclusioni? In realtà, il filosofo americano non nega affatto la possibilità che le macchine possano pensare – egli ritiene (cartesianamente!) che l'uomo sia una macchina e che l'uomo pensi – ma esclude solo che possano farlo sulla base del semplice software che in esse viene fatto girare e ciò perché «gli stati e gli eventi mentali sono letteralmente un prodotto del funzionamento del cervello, mentre il programma non è un prodotto del calcolatore»¹¹. Contro tali conclusioni, Chalmers ha proposto il principio di implementazione, secondo il quale «un sistema fisico *implementa* una computazione quando la struttura causale del sistema rispecchia la struttura formale della computazione»¹²; se è vero, quindi, che i programmi sono puramente sintattici, la loro implementazione fisica sarebbe dotata – invece – di contenuto semantico. Basterebbe individuare la classe di computazio-

¹⁰ J.R. Searle, *La mente è un programma?* cit., pp. 6-8. Searle riprende in gran parte quello che Turing chiama l'argomento della coscienza, presentato da Jefferson nel 1949, cfr. *Calcolatori e intelligenza* cit., p. 67. L'esperimento mentale di Searle è criticato duramente da Hofstadter e Dennet soprattutto per la ragione che esso vorrebbe presentare come normale una situazione in realtà del tutto improponibile, visto l'enorme tempo che sarebbe necessario a un umano per operare con gli ideogrammi cinesi come un computer lavora coi propri simboli formali. In altri termini, Searle giocherebbe sull'ordine di grandezza che rimane incomparabile, e lo fa «poiché altrimenti un lettore scettico si renderebbe conto che la comprensione deve trovarsi quasi tutta in quei miliardi di simboli sulla carta e quasi per nulla, in pratica, nel demone», e cioè nell'umano rinchiuso nella stanza (D.R. Hofstadter e D.C. Dennet, *L'io della mente* cit. p. 362). La percezione e la comprensione, la conoscenza insomma, starebbero al livello dell'intero sistema e non al livello del sé.

¹¹ J.R. Searle, *Menti, cervelli e programmi* cit., p. 357.

¹² D.J. Chalmers, *La mente cosciente* (*The Conscious Mind*, 1996), trad. it. di A. Paternoster e C. Meini, prefazione di M. Di Francesco, Milano, McGraw-Hill, 1999, p. 323.

ni più adatta per replicare con successo la mente umana. A Chalmers si potrebbe tuttavia obiettare che la mente non è una *tabula rasa* anche perché essa è inseparabile dal corpo, il quale ha un corredo di caratteristiche e di potenzialità innate, una *natura* con la quale è indispensabile confrontarsi se non si vuole girare nel vuoto ontologico, nell'eterea mancanza d'attrito che lascia l'indagine ferma al suo punto d'avvio.

È certo vero che le potenzialità degli elaboratori e delle reti informatiche sono notevoli. Alcune delle previsioni formulate solo pochi anni orsono si vanno di fatto realizzando, e spesso oltre le stesse speranze di chi le ipotizzava. Almeno su un punto, però, la scienza dei calcolatori non sembra aver compiuto reali progressi: sulla questione dell'*intelligenza* delle macchine. I calcolatori, infatti, elaborano l'informazione assai più velocemente degli umani, ma non raggiungono la consapevolezza – neppure elementare – di ciò che accade nei loro stessi circuiti, limitandosi a costruire e a computare una serie infinita di uno e di zero. Ora, poiché comprendere implica non solo la padronanza delle *regole* del discorso, il dominio sintattico delle procedure, ma richiede anche la comprensione dei *significati* e quindi una semantica, anche le macchine più potenti sono *ancora* stupide. Lo rimarranno per sempre?

3. Intelligenza Artificiale e antropocentrismo

L'aspirazione di Faust al potere ha molteplici significati e fra questi c'è la tentazione del controllo tecnico sul mondo. Un dominio che non si ferma alla Natura – alle imponenti dighe costruite sul mare ridotto a «feudo» concesso dall'Imperatore per i servizi resigli – e non si interrompe all'*Oggetto* che sta di fronte all'uomo, ma coinvolge totalmente il *Soggetto* e fa dell'antropologia un ambito della tecnologia. Uno degli obiettivi principali di Faust, si potrebbe dire, è la creazione di una forma nuova dell'umano attraverso la produzione di una diversa specie di intelligenza, una *mente artificiale* progettata come il più grande risultato della *mente naturale*. Davanti alla creatura che si sta formando nella fiala, Wagner osserva – presago – che «un cervello che debba

pensare esattamente / lo farà, un giorno, un pensatore»¹³. Come sempre, è Mefistofele – non Faust! – a trarre le rigorose conseguenze di quanto accade, a *capire* il senso degli eventi: «E si finisce che si dipende / dalle creature fatte da noi»¹⁴. L'ambiguo pessimismo di Mefistofele è una tentazione da evitare così come lo è la ὑβρις di Faust. Ma che cosa il futuro ci prospetta realmente? Quali potrebbero essere le notizie dei prossimi anni? Eccone alcune (elaborate da «Scientific American», una rivista rigorosa e dall'impianto concettuale «positivista»): «clonato un essere umano; pronti uteri, cuori e fegati di ricambio; è possibile sentire la realtà virtuale e immergersi in essa; gli abiti su misura diventano alla portata di tutti tramite una perfetta scansione del proprio corpo; trovati dei cibi terapeutici e dei vaccini genetici che permetteranno a tutti di conservare senza fatica e per un tempo assai lungo una muscolatura quasi perfetta...». Queste e altre analoghe novità dovrebbero consentire un allungamento consistente della vita umana e anche un suo miglioramento *qualitativo*.

In ogni caso, le probabili novità radicali che ci attendono nascono tutte dalla «imminente fusione tra biologia e microelettronica, con tutto il suo potenziale di benefici e imprevisti»¹⁵. Fra i primi ci sarebbero le numerose applicazioni terapeutiche rese possibili dalla sintesi fra il corpo umano e una serie di *chip* elettronici in grado di rigenerare tessuti o almeno di consentire di muoversi nello spazio e nella vita con l'ausilio di organi interni ed esterni molto simili a quelli malati o perduti. Sembra che in questi progetti pulsi ancora un'antica tendenza a liberarci dai limiti del corpo umano. In realtà, il settore di ricerca che nel 1986 Jaron Lanier definì «realtà virtuale» è ancora all'inizio del suo sviluppo e ha davanti a sé delle difficoltà non indifferenti se vuole diventare qualcosa di diverso dalla semplice simulazione ludica del mondo fisico. Un gruppo di scienziati che si sta attivamente occupando della possibile integrazione di strutture artificiali e biologiche, allo scopo di curare lesioni di tipo neurologico, ammette comunque che «il cervello e il midollo spinale contengono da 10 a 15 miliardi di cellule ner-

¹³ J. W. Goethe, *Faust*, vv. 6869-6870, trad. it. di F. Fortini, Milano, Mondadori, 1990.

¹⁴ *Ibid.*, vv. 7003-7004.

¹⁵ G. Zorpette e C. Ezzell, in «Futuro bionico. Le Scienze dossier» 4, estate 2000, p. 5.

vose, per di più collegate in maniera complessa, con grande ridondanza. Davanti a una tale mole di dati, il compito di cogliere individualmente e simultaneamente l'attività delle singole cellule nervose e di decifrarne il significato ai fini della percezione o della motilità appare tuttora insormontabile»¹⁶. Se la filosofia ha quindi bisogno di penetrare nelle questioni aperte dall'IA per proseguire nell'indagine da essa avviata sin dall'inizio sulla mente e il suo λόγος, anche l'IA necessita della misura filosofica al fine di chiarire i propri scopi, metodi, concetti e per comprendere che non c'è mai nulla di ingenuamente *neutrale* nella strumentazione tecnica che gli umani approntano a se stessi. Una strumentazione che nell'IA sembra voler assorbire, alla fine, lo stesso inventore, attraverso un'autentica *metamorfosi* della sua natura. L'IA include quindi lo studio sia del biologico che dell'artificiale, sia dell'attuale – frutto della natura oppure dell'intervento umano – sia del potenziale, del probabile, del possibile. L'essere e il nulla, l'ordine e il disordine, il sostanziale e l'accidentale, la misura e l'infinito, il naturale e il prometeico, sono solo alcune delle categorie di una possibile filosofia delle menti artificiali. Un'ontologia dell'IA diventa, in tal modo, un itinerario alle radici stesse della filosofia.

Per un essere umano, vivere significa avere una quantità di informazione che gli permetta di orientarsi nella vita quotidiana, nel tempo e nello spazio e di retroagire rispetto agli stimoli che la realtà continuamente produce e anche impone. In questo senso, non sono poche le analogie fra un essere umano e una macchina. Sia il comando che l'apprendimento possono essere descritti come l'invio di un messaggio che modifica il comportamento di chi lo riceve. Un ordine volto all'esecuzione di un compito può essere quindi impartito a un umano come a un meccanismo, ed esistono macchine così perfezionate da imparare dagli effetti dei loro movimenti in modo da modificare adeguatamente gli atti successivi, al fine del migliore espletamento possibile della propria funzione. L'ambito degli scacchi, ad esempio, non è soltanto ludico e da sempre ha rappresentato un campo cruciale per la messa alla prova delle teorie su una possibile Intelligenza Artificiale e per le realizzazioni pratiche della cibernetica. Esso rappresenta infatti una delle attività

¹⁶ J.U. Meyer, C. Blau e T. Stieglitz, *ibid.*, p. 32.

più formalizzate nelle quali una mente possa esercitarsi. Ebbene, nel 1950 Wiener scriveva che «non è difficile costruire una macchina che, osservando meccanicamente le regole, giochi una partita a scacchi di livello assai scadente ma è impossibile costruire una macchina che si comporti come un eccellente giocatore di scacchi»¹⁷ e quasi tutti gli studiosi erano convinti che mai una macchina avrebbe potuto sconfiggere un Gran Maestro degli scacchi, poiché si tratta di un gioco nel quale la pura potenza tattica del calcolo – la forza bruta della infinita serie di zero e di uno che un computer può mettere in fila – non basta e dove invece è necessaria una comprensione strategica del gioco e del suo contesto. E tuttavia Kasparov – all'epoca campione del mondo – nel 1997 fu sconfitto da *Deep Blue*, con una modalità che fece dire allo sfidante umano che la macchina aveva dato – per quanto incredibile potesse apparire – «segni di intelligenza». Si tratta di un evento certamente rilevante, anche se rappresenta solo un primo passo nell'evoluzione dei computer verso l'intelligenza.

In tale cammino, una delle intuizioni fondamentali della cibernetica è stata l'aver colto l'importanza della comunicazione di *bit* rispetto alla trasmissione degli *atomi*. Ben prima di Negroponte, Wiener ha teorizzato la possibilità della «trasmissione di messaggi di comunicazione che non comportano lo spostamento di una particella di materia da un capo all'altro della linea. Abbiamo così due tipi di comunicazione: e cioè il trasporto materiale e il semplice trasporto di informazione»¹⁸. È già realtà affermata e quotidiana, e da molto tempo ormai, il trasferimento nello spazio di parole, immagini, suoni, in tempo reale e con modalità che nulla fanno perdere del documento-corpo originario. La cibernetica rappresenta pertanto un ambito epistemo-

¹⁷ N. Wiener, *Introduzione alla cibernetica (The Human Use of Human Beings, 1950)*, trad. it. di F. Ciafaloni, Torino, Bollati Boringhieri, 2001, p. 219. Sui precedenti nel mondo greco dei progetti di «robotica» – da Ctesibio al teatro degli automi di Erone di Alessandria – si veda l'ampia ricostruzione della scienza ellenistica elaborata da L. Russo, *La rivoluzione dimenticata. Il pensiero scientifico greco e la scienza moderna*, nuova ed., Milano, Feltrinelli, 2001, in particolare le pp. 152-159, con la conclusione critica secondo la quale «molti degli apparecchi di Erone potrebbero essere interpretati come sottoprodotti della tecnologia ellenistica (inizialmente sviluppata per altri scopi) che proprio grazie alla loro natura ludica erano riusciti a sopravvivere e a svilupparsi nelle nuove condizioni dell'età imperiale» (p. 159).

¹⁸ *Ibid.*, p. 121.

logico dalle profonde implicazioni metafisiche, nel quale concetti come *tempo*, *spazio*, *enti*, *comunicazione*, *entropia* acquistano significati peculiari, in parte coerenti con immagini della realtà fisica come quella kantiana o einsteiniana, ma anche in radicale discontinuità con una percezione dello spazio-tempo, e quindi del mondo, legata in modo indissolubile alla materia.

Non a caso all'inizio degli anni cinquanta la discussione sulla nuova scienza creata da Wiener fu molto viva anche in ambito filosofico. Nel 1953 Hans Jonas, ad esempio, durante una conferenza tenuta presso la New York School for Social Research affrontava uno dei temi centrali dell'epistemologia cibernetica, il teleologismo delle macchine e la sua analogia con quello degli umani. Jonas coglie con chiarezza le ambizioni della cibernetica proprio nel suo impianto aristotelico, volto alla comprensione dei fini e consapevole della loro centralità ontologica. Di più, come l'aristotelismo la cibernetica rappresenta una dottrina unificata per la rappresentazione dell'intera realtà a partire dai concetti di informazione, scopo, comunicazione. Ma proprio in questa ambizione Jonas coglie un pericolo per la filosofia e cioè per una conoscenza integrale della realtà. Sin dall'inizio, infatti, «la cibernetica ha cercato di conquistare la filosofia. Fin dall'inizio ha preteso lo *status* di una teoria unificata di meccanicismo, organicismo, sistema nervoso, società, cultura e mente»¹⁹. Per Jonas la pretesa della cibernetica di comprendere coi suoi strumenti e con le sue categorie l'intero mondo animale, artificiale e umano si scontra, quindi, con alcuni ostacoli insuperabili e con dei limiti di fondo.

Anzitutto, anche la cibernetica – come il comportamentismo e la biologia meccanicistica – subisce la contraddizione logica di escludere dai propri contenuti e postulati proprio coloro, gli scienziati, che indagano attraverso i suoi metodi. In secondo luogo, «l'intera dottrina cibernetica del comportamento teleologico è riducibile ad una confusione tra “servire a uno scopo” e “avere uno scopo” ed in modo particolare alla confusione tra “perseguire uno scopo” e averne uno»²⁰. Le macchi-

¹⁹ H. Jonas, *La cibernetica e lo scopo: una critica*, a cura di M. Campanelli, Pisa, ETS, 1999, p. 58.

²⁰ *Ibid.*, p. 53.

ne, infatti, possono anche modificare i propri movimenti per effetto di una informazione che esse stesse riescono a elaborare, e ciò tramite lo strumento della retroazione o *feed-back*, ma rimangono poi totalmente indifferenti alla riuscita o meno dell'azione finalizzata a uno scopo. L'essere umano, invece, avrebbe già in partenza un fine che è frutto di una serie complessa e articolata di *motivazioni* dei generi più diversi e, di fronte alla riuscita o al fallimento della propria azione, reagisce tramite sentimenti di soddisfazione o di amarezza, rabbia, frustrazione. In altri termini, e questo sarebbe il terzo limite della cibernetica, anche la scienza dell'informazione e della comunicazione rappresenta una forma di riduzionismo, in quanto «le cose viventi sono creature del bisogno» e invece «il modello cibernetico riduce la natura animale a due termini: facoltà di sentire e di movimento, quando di fatto essa è costituita da una triade: percezione, facoltà di movimento ed emozione»²¹. E tuttavia bisogna anche riconoscere che Wiener aveva escluso in maniera decisa la semplice trasposizione degli schemi di informazione-comportamento delle macchine al comportamento e alle motivazioni umane, affermando esplicitamente che il linguaggio umano si distingue da qualunque forma non umana di comunicazione, compresa quella degli scimpanzé che pure sono i più vicini a noi geneticamente; aveva ripetuto che una macchina non potrà mai comprendere il gioco degli scacchi come fa un umano, e quindi non potrà mai *giocare* allo stesso modo e, soprattutto, aveva sostenuto che per quanto possano venir perfezionate, alle macchine non si debba cedere nulla della *responsabilità* dell'azione, la quale rimane tutta e integralmente una prerogativa umana.

Dove invece la critica di Jonas coglie una delle effettive caratteristiche della cibernetica è nel concetto di entropia. Se lo «scopo» di una macchina o di un organismo è il raggiungimento «di una condizione finale in cui l'oggetto di cui si studia il comportamento raggiunge una correlazione definita nel tempo o nello spazio rispetto a un altro oggetto o evento»²², il fine ultimo di ogni azione e dell'intera esi-

²¹ *Ibid.*, p. 59.

²² A. Rosenblueth, N. Wiener e J. Bigelow, *Comportamento, scopo e teleologia* (1943), in *La filosofia degli automi. Origini dell'intelligenza artificiale*, a cura di V. Somenzi e R. Cordeschi, Torino, Boringhieri, 1986, p. 68.

stenza sarebbe la condizione di entropia assoluta, la morte. E tuttavia, «un meccanismo di *feed-back* può funzionare o stare fermo: in entrambi i casi la macchina esiste. L'organismo deve continuare a funzionare, perché in questo consiste la sua esistenza – che è provvisoria – e, minacciato dalla sua fine, si afferma attraverso la sua esistenza. Non c'è nelle macchine qualcosa di analogo all'istinto di autoconservazione, ma soltanto la sua antitesi, cioè l'entropia finale della morte»²³. Ma le cose stanno davvero o *soltanto* in questo modo? La tensione estrema che è la vita non sembra avere come scopo ultimo il suo cessare? La precarietà dell'esistere, la sua natura quasi casuale, frammentaria e finita, non condiziona il vivere sociale e il corpo singolo nella inevitabilità della distruzione che in un moto incessante sostituisce vite a vite, esseri a esseri senza che questo abbia un senso fuori di un eterno e gratuito avvicendamento? Il metabolismo²⁴ come forma base dell'esistenza non rappresenta anch'esso un circolo di fattualità-nutrimiento-replica che non sembra avere altro fine se non se stesso? Inserire in questo tipo di indagine, come fa Jonas, concetti quali *consapevolezza*, *libero arbitrio*, *ricerca del bene* e *somnum bonum* non comporta l'introduzione di elementi moralistici, valutativi, antropomorfici? Se la cibernetica è il luogo in cui la filosofia finisce non solo nel senso che in essa si *conclude* ma anche e soprattutto nel senso che in essa si *trasforma*, «la fine della filosofia è contrassegnata dalla dissoluzione delle sue discipline in scienze autonome, la cui unificazione sotto nuova forma si profila nella cibernetica»²⁵. I processi di informazione, controllo e retroazione trovano nell'Intelligenza Artificiale il loro elemento unificante, ma questo significa anche che la cibernetica e l'IA costituiscono l'orizzonte del pensiero contemporaneo, quell'orizzonte nel quale bisogna *stare* e che bisogna *comprendere* affinché per il pensiero ci sia ancora una possibilità.

²³ H. Jonas, *La cibernetica e lo scopo...* cit., p. 60.

²⁴ Oggetto delle indagini jonasiene di *Organismo e libertà. Verso una biologia filosofica*, a cura di P. Becchi, Torino, Einaudi, 1999. Il metabolismo sarebbe, secondo Jonas, già qualcosa di qualitativamente diverso rispetto all'esistenza minerale possedendo in sé il germe dell'autonomia dal semplice causalismo meccanico. I gradi successivi dell'esistenza animale consistono nella motilità, percezione e sentimento, i quali giungono alla loro più chiara manifestazione nell'essere umano.

²⁵ M. Heidegger, *Filosofia e cibernetica (Zur Frage nach der Bestimmung der Sache des Denkens)*, 1984), nuova edizione riveduta e ampliata a cura di A. Fabris, Pisa, ETS, 1997, p. 33.

4. Intelligenza Artificiale e problemi dell'apprendimento

La questione di una possibile Intelligenza Artificiale mostra pertanto tutto il suo senso e la sua centralità epistemologica all'interno di una più ampia prospettiva antropologica che ponga in modo esplicito la domanda sull'identità e la struttura di ciò che definiamo *umano* e su quanto di ciò che lo caratterizza sia: a) immodificabile; b) costitutivo; c) esclusivo.

La specie umana così come la vediamo e siamo non rappresenta il vertice dell'essere e del tempo. Se il risultato che noi stessi costituiamo è probabilmente il più raffinato e complesso che si dia, questo non implica che esso sia anche l'*ultimo*. L'indagine etologica di Lorenz afferma esplicitamente la natura provvisoria dell'umano: Lorenz ricorda infatti che l'attuale specie umana «è soltanto un effimero anello nella catena delle forme viventi. Ci sono buone ragioni per pensare che [essa] sia soltanto un gradino nella scala che porterà a un essere *realmente umano*. Se non altro, è possibile sperarlo»²⁶. E quindi, «molto lontano dal vedere nell'uomo la definitiva e insuperabile immagine di Dio, affermo con più umiltà e, come credo, con maggior rispetto per la creazione e le sue inesauribili possibilità: il tanto ricercato anello intermedio fra l'animale e l'uomo veramente umano, *siamo noi!*»²⁷. Sarebbe pertanto immotivata ogni riflessione sulla natura umana che assuma ciò che siamo nel presente come un modello inoltrepassabile, come un paradigma dato per sempre, come un valore indiscutibile. Veniamo certamente da molto lontano e altrettanto lungo è il cammino che possiamo ancora percorrere. Eibl-Eibesfeldt arriva a ritenere probabile «che cambiamenti culturali dello stile di vita possano indurre in futuro anche cambiamenti genetici; a favore di questa ipotesi vi sono già buoni indizi»²⁸.

²⁶ K. Lorenz, *Il declino dell'uomo (Der Abbau des Menschlichen)*, 1983), trad. it. di A. Casalegno, Milano, Mondadori, 1995, p. 229.

²⁷ Id., *L'aggressività (Die sogennante Böse: zur Naturgeschichte der Aggression)*, 1963), trad. it. di E. Bolla, introduzione di G. Celli, Milano, Mondadori, 1990, p. 300.

²⁸ I. Eibl-Eibesfeldt, *Etologia umana. Le basi biologiche e culturali del comportamento (Die Biologie des menschlichen Verhaltens Grundriss der Humanethologie)*, München, R.Piper GmbH e Co. KG, 1984), edizione italiana a cura di R. Brizzi e F. Scapini (con gli aggiornamenti dell'autore per l'edizione USA 1989), Torino, Bollati Boringheri, 1993, p. 12. Sulle più generali implicazioni antropologiche della questione che stiamo affrontando, si vedano di A.G. Biuso, *Antropo-*

Se il pensare è una caratteristica specifica della nostra specie, può il pensiero nascere e operare fuori da un legame strettissimo con il corpo? I «qualia» – le sensazioni che *si provano* a essere un determinato ente e non un altro –, le esperienze fenomeniche, individuali, qualitative possono essere separate dal legame con l'organico, coi sensi, con il biologico? Gli studiosi che criticano il programma forte dell'IA fondano il loro rifiuto proprio su questo punto, sulla necessità per qualunque mente di possedere un corpo che trasformi il linguaggio formale della programmazione in una fonte di conoscenza reale, in una esperienza anche socialmente incarnata. Dato però che «nel computer non si trova incorporata alcuna *pre-comprensione* di come sia organizzato il nostro mondo e di come in esso ci si debba aggirare», nessun computer – per quanto potente e adeguatamente programmato – potrà mai davvero pensare²⁹. Uno dei problemi e dei limiti di fondo del progetto IA sembra quindi l'assenza di un *corpo organico*, delle sue esperienze, della crescita come travaglio e gloria della corporeità. Senza il corpo non ci sarebbe pensiero ma solo calcolo.

Alla interpretazione puramente simbolico-sintattica della mente, la *embodied cognitive science* sostituisce una prospettiva per la quale mente, cervello e mondo sono parti di un'unica struttura. La mente non sarebbe quindi delimitata dai confini della pelle o della scatola cranica. Essa non solo sarebbe diffusa in tutto il corpo – come il neurologo Antonio Damasio e ancor prima di lui Arnold Gehlen hanno ipotizzato –, ma ciò che chiamiamo mente sarebbe in realtà l'insieme del cervello umano più l'ambiente col quale il cervello intrattiene continue relazioni di comprensione, adattamento, osmosi³⁰. In sintesi: «i nostri non sono cervelli

logia e Filosofia, Napoli, Guida, 2000 e *Oltre il dualismo natura/cultura. L'integrazione possibile di etologia e antropologia*, in «Discipline Filosofiche» XII/1, Macerata, Quodlibet, 2002, pp. 229-254

²⁹ H.L. Dreyfus, «Una risposta ai miei critici» in T.W. Bynum e J.H. Moor, *La fenice digitale. Come i computer stanno cambiando la filosofia (The Digital Phoenix. How Computers are Changing Philosophy)*, 1998), Milano, Apogeo, 2000, p. 215.

³⁰ Di Damasio si vedano *L'errore di Cartesio (Descartes' Error)*, 1994), trad. it. di F. Mancuso, Milano, Adelphi, 1995 e il recente *Looking for Spinoza*, New York-London, Harcourt, 2003, dove si sostiene che «the images that flow in the mind are reflections of the interaction between the organism and the environment, reflections of how the brain's reaction to the environment affects the body» (p. 206). Nella sua opera principale Gehlen rileva che «con grande facilità si commette l'errore generale di localizzare l'intelligenza dell'uomo nella sua testa, trascurando la grande ra-

di spiriti disincarnati convenientemente incollati dentro gusci corporei ambulanti di carne e sangue. Piuttosto essi sono *essenzialmente* i cervelli di agenti incarnati, capaci di creare e sfruttare strutture nel mondo»³¹. Una Intelligenza Artificiale forte sarebbe quindi impossibile proprio perché i computer sembrano abilissimi nel riprodurre le funzioni alte della mente, quelle appunto logico-simboliche, ma del tutto incapaci di percepire il senso comune, di elaborare il linguaggio naturale, di utilizzare l'intelligenza senso-motoria. Questo vero e proprio paradosso dell'Intelligenza Artificiale segnerebbe secondo Dreyfus lo scacco definitivo del razionalismo socratico e la rivalsa di Eutifrone rispetto a Socrate³². Alle tesi di Dreyfus si può tuttavia rispondere che la corporeità – naturale dell'essere umano o artificiale dei sensori applicati a una macchina – è una condizione necessaria ma non sufficiente dell'intelligenza.

Notevole è l'impatto di queste prospettive sull'insegnamento. Insieme alle grandi potenzialità, non bisogna nascondere i possibili rischi. La fiducia ingenua e acritica nell'uso dei computer in ambito educativo si rivela uno dei tanti *miti d'oggi*: «i calcolatori non sono una bacchetta magica per migliorare la scuola. Se vengono usate per aiutare gli studenti a imparare a memoria, le macchine sono ben poco utili. Consentendo agli studenti di interagire con ricche risorse intellettuali, tra cui insegnanti, biblioteche e musei lontani, l'infrastruttura può però diventare un potente alleato nel settore dell'istruzione»³³. Il burocratico ottimismo dei ministeri ha riempito, ad esempio, le scuole italiane di personal computer presto obsoleti, il cui utilizzo didattico è ancora un mistero per la gran parte dei docenti e che si sono quindi ridotti a uno dei tanti elementi dell'arredo scolastico... Ben altre sembrano, invece, le potenzialità dei computer e della Rete come strumento di scambio di informazioni nell'ambito della ricerca, settore nel quale Internet è nata

gione del corpo, lasciato ai fisiologi e a altri specialisti» (*L'Uomo. La sua natura e il suo posto nel mondo* [Der Mensch. Seine Natur und seine Stellung in der Welt, 1978, I ed. 1940], trad. it. di C. Mainoldi, Milano, Feltrinelli, 1990, p. 397).

³¹ A. Clark, *Dare corpo alla mente* (*Being There*, 1997), trad. di S. Levi, Milano, McGraw-Hill, 1999, p. 194.

³² Si veda il saggio di G. Stelli, *La rivincita di Eutifrone: didattica postmoderna e Intelligenza Artificiale*, in «Punti Critici» 5/6, dicembre 2001, pp. 13-62.

³³ M. Dertouzos, in «Reti informatiche. Quaderni Le Scienze» 95, a cura di P. Capiluppi, aprile 1997, p. 25.

e che la va sempre più trasformando, come scrive efficacemente Gary Stix, nella «più grande lavagna del mondo»³⁴. Nella interazione tra informatica e apprendimento è opportuno distinguere tra i semplici «*chioschi informativi* (testi, antologie di testi, enciclopedie, lezioni)» dal livello di interattività piuttosto ridotto e i «*CBL (Computer-Based Learning)*, applicazioni costruite avendo come obiettivo principale l'*interattività*», da non confondersi – in ogni caso – con la semplice *multimedialità*; «il computer [...] va considerato, a livello, didattico, essenzialmente come uno strumento e nessuno strumento può sostituire l'insegnante e la sua opera»³⁵.

Stelli ha discusso in modo convincente le forme e i limiti di una didattica postmoderna, la quale ha come principale obiettivo la deconcettualizzazione dell'insegnamento e dell'apprendimento a favore delle esperienze *primarie ma ritenute esclusive dell'umano* di tipo senso-motorio, manipolativo, iconico. Derivano da qui le posizioni e le proposte che vedono nella multimedialità e ipermedialità una svolta decisiva in ambito sia didattico che epistemologico. La corporeità è senza dubbio uno degli elementi caratterizzanti e imprescindibili dell'essere umano come di qualunque ente che viva in uno spazio-tempo fisico. Non per questo, tuttavia, è legittima l'identificazione della *mente* con il *cervello* e cioè con il sostrato materiale che la elabora. Attribuire, poi, all'esperienza senso-motoria un valore identificante e capace di distinguere l'ente umano dalle macchine pensanti può risultare utile ai fini della polemica contro le ipotesi forti di Intelligenza Artificiale, ma comporta anche l'inevitabile sottovalutazione delle componenti formali, simboliche, concettuali non solo e non tanto del linguaggio umano, ma dello stesso posto dell'uomo nel mondo. Sono molto significative le ricadute didattiche di questo riduzionismo: attribuire un primato epistemologico e formativo a tutto ciò che *non* è concetto e pensiero astratto non ha comportato la liberazione dal dominio dell'artificiale, ma esattamente il contrario, l'elevazione del computer a mito, la sua trasformazione da strumento didattico fra gli altri in macchina primaria, unica ed esclusi-

³⁴ *Ibid.*, p. 66.

³⁵ G. Stelli e D. Lanari, *Modelli di insegnamento della filosofia. Modello teoretico, modello storico, filosofia al computer*, Roma, Armando, 2001, pp. 81-82.

va di apprendimento. La «inevitabile» sostituzione dei libri con apparati multimediali – preconizzata dai fautori della didattica deconcettualizzata – è sempre più smentita dai fatti. E tuttavia mentre simili previsioni si mostrano quanto meno azzardate, la deconcettualizzazione dell'insegnamento si impone nel modo soft e indiretto dell'abbassamento dei livelli di apprendimento richiesti a chi frequenta le scuole³⁶.

L'*Interconnected network* è un dispositivo assai potente di comunicazione, ma non è altro che il più avanzato degli strumenti – pietra, papiro, carta... – che rendono possibile quello scambio infinito fra menti in cui consiste l'edificio della socialità e della cultura. Il *Web*, la ragnatela di dati, di testi, di immagini, di bit che corre sui supporti telematici è uno dei modi coi quali opera e comunica la specie umana. Le sue potenzialità sono enormi ma si tratta pur sempre di abilità strumentali. La forma della comunicazione rimane totalmente dipendente dal suo contenuto. In questo senso, entusiasmi e timori – nei confronti del nuovo mezzo – sono entrambi ingiustificati. Nell'era di Internet, come in quella dell'Atene classica, delle Accademie rinascimentali o delle università moderne, la comunicazione più efficace fra allievo e maestro, la più vera delle strategie conoscitive rimane quella socratica: «in realtà, il docente può fare tutto e il contrario di tutto, ma alla fine se lo studente impara è perché a capire un determinato concetto c'è arrivato lui, proprio lui, nessun altro lo ha potuto fare al posto suo»³⁷.

³⁶ Utile a chiarire questi temi è un saggio di Sergio Guarente dal titolo «Filosofia e Intelligenza Artificiale: dal dibattito teorico alle implicazioni metodologico-didattiche» in *Lo spazio della filosofia. Riflessioni e proposte didattiche*, a cura di P. Sensi e G. Stelli, Perugia, GESP, 2002. In esso l'Autore traccia una sintesi molto chiara dei rapporti tra filosofia e IA, a partire dalla convinzione che si tratta di due ambiti fra di loro strettamente collegati. L'IA, infatti, non è una disciplina ingegneristico-tecnica, ma si fonda su una ben precisa concezione del pensare umano e il suo sviluppo implica importanti effetti sul piano epistemologico. Alcune delle conseguenze didattiche di questo dibattito fra i sostenitori e i critici dell'IA sono analizzate da Giovanni Stelli, che in questo stesso volume presenta una sua proposta di ripensamento del modello storico della filosofia in funzione antistoricistica, convinto che sia un errore confondere questi due modelli per poi negare legittimità all'insegnamento della storia della filosofia (p. 42). Contro le eccessive speranze che molti tecnologi della didattica ripongono negli strumenti informatici e multimediali, Stelli e Guarente convergono poi sul fatto che in ogni caso occorra «non perdere di vista nell'azione didattica le abilità logico-astratte proprie del pensiero scientifico e storicamente legate all'alfabetizzazione tipografica» (p. 160). Un'esigenza che ci trova pienamente concordi.

³⁷ D. Massaro e A. Grotti, *Il filo di Sofia. Etica, comunicazione e strategie conoscitive nell'epoca di Internet*, Torino, Bollati Boringhieri, 2000, p. 193.

La macchina da sola non produce nessun miracolo virtuale, nessuna rivoluzione epistemologica, proprio perché non è possibile un insegnamento separato dalla concreta e viva soggettività del docente. Ciò non significa, naturalmente, che strutture scolastiche e programmi formativi non debbano essere pensati attentamente, ma solo che debbano essere pensati in funzione del vivo rapporto *educativo* con gli allievi e delle rigorose competenze *disciplinari* di chi sta in cattedra. Nessuna riforma e nessuna novità metodologica produrranno degli effetti positivi, se prima di tutto chi insegna non conosce a fondo la propria disciplina. Nessuna riforma che voglia avere successo può permettersi di trascurare le competenze e l'esperienza professionale di maestri e professori. La tecnologia del Web da sola non dà valore aggiunto all'insegnamento, se questo non ha alla sua base una pedagogia matura e consapevole. Dall'equivoco – ingenuo ma diffuso – della autonomia del mezzo rispetto al fine nasce quella realtà ben sintetizzata da Barnette: «le attività dell'insegnamento della filosofia nel cyberspazio ispirano e gratificano così come spesso deludono e creano dubbi»³⁸. Chi sta ogni giorno a scuola e cerca di utilizzare nella propria didattica i laboratori informatici sa che le cose stanno proprio in questo modo. Nessuna chiusura quindi nei confronti dell'e.learning ma anche nessun entusiasmo acritico.

L'evoluzione degli strumenti informatici – ad esempio, i libri elettronici – ha infatti reso ancora più evidente la differenza ontologica fra i contenuti culturali e mentali, il *significato*, e i supporti materiali che li veicolano, il *significante*. Un libro, infatti, che cos'è? Il testo o il supporto (papiro, pergamena, carta, schermo LCD) che lo veicola³⁹? Attraverso la Rete, del libro rimane l'essenziale del mondo 3 e cioè il contenuto informativo che lo costituisce, mentre si modifica la maniera in cui tale contenuto viene distribuito, il mondo 1: non più una libreria con degli scaffali ma una serie di impulsi elettronici.

Le forme della comunicazione mutano quindi nel tempo e nello spazio, ma si radicano tutte nella capacità di elaborare segni che stiano al posto delle cose e segni che indichino altri segni. Il linguaggio umano

³⁸ R. Barnette, «L'insegnamento della filosofia nel cyberspazio», in *La fenice digitale* cit., p. 367.

³⁹ Sui rapporti fra la Rete e i libri, si veda A.G. Biuso, *HyperNietzsche. La Rete, Nietzsche, il futuro dei libri*, in «La Rivista dei Libri», maggio 2001, pp. 29-31.

non è solo *denotativo* ma è soprattutto *significativo*. Una parola acquista la sua completezza soltanto in un contesto che le dia un preciso significato. La formalizzazione dei simboli produce comprensione, scambio e azione in quanto i segni rinviano al gioco reciproco fra gli enti e le menti. Comunicare è un'attività che si serve di strumenti fisici in vista di uno scopo immateriale. Attraverso dei suoni o dei simboli grafici è il pensiero che transita; una mente individuale entra in contatto *diretto* con un'altra mente. A cambiare sono le modalità di questo passaggio, non la sua natura. È anche per questo che è possibile comunicare con individui il cui corpo si è dissolto da tempo, ma i pensieri dei quali sono più vivi che mai. La *scrittura* rappresenta uno degli strumenti essenziali di tale relazione: non importa il supporto sul quale si imprimono i segni, conta il loro significato che è immateriale e quindi resistente al tempo. In qualche modo, è l'intero mondo della scrittura, è tutto il linguaggio umano a essere di per sé *virtuale*, se si intende con questo termine la dimensione non fisica di ciò che gli enunciati trasmettono materialmente. Gli enti del mondo 3 possono essere espressi in una pluralità di forme del mondo 1 e proprio per questo rimangono irriducibili a tale mondo. Il pensiero che si manifesta nell'atto dello scrivere dipende certamente da tale azione, ma la potenzialità di pensare – la forma a priori della mente – precede sia i singoli pensieri che i vari modi della loro espressione e produce un mondo che può essere riversato poi su una molteplicità di supporti dai quali rimane sostanzialmente indipendente.

5. Conclusione: dai Greci all'*Artificial Intelligence*

Per degli esseri umani è naturale concepire il pensiero come una serie di dati posti l'uno dopo l'altro e in collegamento reciproco. Una formulazione chiarissima delle possibilità insite in questo modo di intendere la conoscenza è fornita da Descartes come significativo commento alle quattro regole del metodo: «quelle lunghe catene di ragionamenti, tutte semplici e facili, di cui i geometri sono soliti servirsi per pervenire alle loro più difficili dimostrazioni, mi avevano dato occasione di immaginare che tutte le cose suscettibili di cadere sotto la conoscenza umana si seguono l'un l'altra nello stesso modo e che, posto sol-

tanto che ci si astenga dall'accettarne qualcuna per vera mentre non lo è e che si osservi sempre l'ordine necessario per dedurle l'una dall'altra, non ve ne possono essere di così lontane cui alla fine non si pervenga, né di così nascoste che alla fine non si scoprono»⁴⁰. L'IA nacque come applicazione di tale modo di elaborazione dei dati ai circuiti elettronici. Una serie di pacchetti di informazione venivano trasmessi alle macchine l'uno dopo l'altro, in sequenza seriale. La velocità con la quale i computer elaboravano questi dati era già grande all'inizio e col tempo è andata crescendo, con l'evolversi dei processori, a valori davvero straordinari. E tuttavia la pura forza bruta del calcolo non ha generato una vera Intelligenza Artificiale. Né sembra che da un'ulteriore crescita essa possa venire realmente prodotta. Alla metafora cartesiana seriale della lunga catena di ragionamenti va forse preferita quella galileiana di un grandissimo numero di formiche che operando in parallelo «strascicherebbe per terra una nave carica di grano» poiché «mentre una resistenza non sia infinita, può dalla moltitudine di minimissime forze esser superata»⁴¹.

È quindi di primaria importanza comprendere che le menti umane e i computer *ragionano* in modo differente poiché sono ambienti differenti. Ne segue che se le macchine saranno davvero in grado di elaborare pensieri lo faranno in un modo diverso da come ci saremmo aspettati all'inizio e in un modo forse difficilmente comprensibile a noi. Postulare che l'intelligenza debba somigliare in ogni caso a quella umana, pena il non essere intelligenza, è probabilmente un residuo di quell'atteggiamento antropomorfo e antropocentrico che già Senofane aveva colto e che si esprime, ad esempio, nella pretesa di trovare nell'universo forme *aliene* di vita che non siano però troppo aliene da ciò che l'umanità ha inteso come vita a partire dalla vita del suo stesso corpo. Così, si costruiscono radiotelescopi in grado di captare segnali lontani, provenienti anche dall'universo profondo, ma si presuppone che tali segnali siano codificati con la regolarità di pacchetti seriali costruiti e ordinati secondo le strutture tipiche delle trasmissioni radio terrestri.

⁴⁰ R. Descartes, *Discorso sul metodo*, in «Opere filosofiche» a cura di B. Widmar, Torino, Utet, 1981, p. 145.

⁴¹ G. Galilei, *Discorsi e dimostrazioni matematiche intorno a due nuove scienze*, Torino, Boringhieri, 1958, p. 32.

Si tratta più probabilmente, se ci sono o se sono possibili, di forme di pensiero e di vita differenti dalle nostre. Avrebbe poco senso stabilire delle gerarchie fra diversi – i terrestri e gli alieni, i computer e gli umani – e sarebbe più ragionevole accettare questa differenza come un esempio della ricchezza dell'essere rispetto alle categorie nelle quali cerchiamo di rinchiuderlo. Non dovremmo, inoltre, mai dimenticare che i computer che abbiamo inventato e che andiamo costruendo – sempre più veloci, affidabili, autonomi – costituiscono un'espressione dell'intelligenza e della identità umane, le quali potrebbero in tal modo transitare dalla forma *biologica* dell'evoluzione a una possibile forma *computazionale*, intesa come l'amplificazione dell'intelligenza umana attraverso l'integrazione con le macchine, in particolare coi computer, mettendo «direttamente in contatto le onde elettroencefaliche con i vari software»⁴².

In questa chiave, assumerebbe un diverso significato anche il fondamentale problema dei *qualia*. Come abbiamo già accennato, con questo termine si esprime una duplice esigenza. La prima fa riferimento al radicarsi del pensiero umano nella struttura del corpo; una struttura che abita e vive nello spazio-tempo multidimensionale, dal contatto col quale deriva la vera peculiarità della mente e la sua differenza *strutturale* rispetto alla capacità di calcolo dei circuiti elettronici, i quali ignorano l'attrito costituito dalla vita, la resistenza – l'hegeliano «travaglio del negativo» – che consente al pensiero umano di crescere, trasformarsi, adattarsi e creare. La seconda esigenza fa riferimento alla dimensione creativa del pensiero che sarebbe preclusa a qualunque forma possibile di IA proprio perché il semplice calcolo, per quanto spinto a livelli attualmente impensabili, non produrrebbe la *coscienza* del calcolare, la passione dell'apprendere, la unicità dell'imparare umano non solo per prove ed errori ma anche – e soprattutto – per scarti e invenzioni. Se accettiamo, però, la possibilità – solo la possibilità – che il

⁴² R. Marchesini, *Post-human. Verso nuovi modelli di esistenza*, Torino, Bollati Boringhieri, 2002, p. 455. Si veda, ad esempio, di A. Oliverio, *Una protesi per la memoria*, in «Mente & cervello» 1/4 (luglio-agosto 2003): «gli sviluppi della bioingegneria hanno permesso di costruire protesi estremamente sofisticate, basate su interfacce tra nervi, muscoli e servomeccanismi», sviluppi che stanno rendendo possibile l'implementazione di un'apparecchiatura che svolga la funzione dell'ippocampo, fondamentale per il consolidamento e la conservazione dei ricordi; «se i risultati saranno positivi, si aprirà una nuova via per riparare alcuni danni del cervello umano e i confini tra naturale e artificiale saranno sempre più tenui» (p. 93).

pensiero non sia o non rimanga una prerogativa esclusiva degli umani, dobbiamo accettare la prospettiva che di *altra forma di pensiero* possa trattarsi rispetto a quella a noi familiare. Si aprono, da qui, altre due possibilità. La prima, ponendosi nel solco della filosofia hegeliana della storia, accetta che lo Spirito si esprima in forme non antropiche e che l'umanità si prepari quindi a consegnare il testimone del pensiero a degli enti nati da lei ma ormai da essa autonomi. È l'implicito di una formula come la seguente: «la prima macchina ultra-intelligente è l'ultima invenzione che l'uomo ha bisogno di costruire»⁴³. La seconda possibilità accetta in pieno, e quindi senza paure per i nostri destini, le potenzialità evolutive offerte dalla computazione e si dispone a una collaborazione-simbiosi tra le menti umane e le macchine da esse inventate. In questo modo, il tempo e la storia assisterebbero a una evoluzione dell'umano – con tutto il condizionamento ma anche la ricchezza della sua *corporeità* – verso lo Spirito, non più nelle tradizionali espressioni religiose, ma in direzione di forme altrettanto immateriali perché radicate nella potenza del numerico/digitale. Potenza delle matematiche la quale fu una volta il principio della metafisica pitagorica e platonica, a loro volta fondamento del pensiero europeo.

⁴³ V. Vinge, *Technological singularity*, in «Whole Earth Review», inverno 1993, p. 90.