



IL CORPO LA MENTE

ROBOT E UOMINI NEL FUTURO
DELL'INTELLIGENZA ARTIFICIALE

ROBERTO CINGOLANI

Se consideriamo il nesso corpo-mente, scopriamo che esso è difficilmente eguagliabile dalla realizzazione di un computer connesso ad attuatori di movimento e a sensori. Grazie alla sua lunga e complessa evoluzione, oggi il cervello umano funziona in modo combinato e contemporaneo con il corpo. Attualmente è impossibile trasferire nelle macchine le sinergie mente-attuazione tipiche dell'uomo, perché l'intelligenza elettronica e il corpo mecatronico operano con meccanismi diversi da quelli biologici. Quale sarà l'evoluzione dell'Intelligenza Artificiale, quali sono le prospettive e i limiti di questo sviluppo, anche in relazione alle incognite di tipo etico che si intravedono?

Nel corso della storia della tecnologia, ricercatori e scienziati hanno lavorato duramente per raggiungere l'obiettivo di creare robot sempre più simili all'uomo e intelligenze sempre più affini a quella umana. L'espressione In-

telligenza Artificiale (AI) è stata coniata nel 1956 dal matematico americano John McCarthy; da allora, man mano che i risultati tecnici e scientifici sono progrediti, scienziati e filosofi hanno iniziato a riflettere con passione e senso critico su quanto si possa parlare di «intelligenza» riferendosi alle macchine, e quanta analogia ci possa essere fra la macchina e l'uomo. Negli anni Sessanta sono pubblicati i primi saggi sull'argomento: con lungimiranza si individuano molti ostacoli che la ricerca sulla AI avrebbe incontrato negli anni successivi, soprattutto a causa della forte differenza fra i risultati cognitivi ottenuti dalle macchine e quelli tipici degli esseri umani.



Nel 1969 Marvin Minsky e Seymour Papert, nel volume *Perceptrons*, evidenziano i limiti delle prime reti neurali artificiali da loro realizzate¹. Data l'autorevolezza degli autori, la pubblicazione spegne l'entusiasmo creatosi attorno a questo tema e determina una consistente riduzione dell'interesse scientifico ed economico nei confronti della AI, uno stallo che si protrarrà per circa un ventennio. Negli anni Novanta, grazie all'accelerazione delle tecnologie elettroniche e all'aumento delle prestazioni dei computer, la ricerca nel campo della AI riprende con maggior vigore: ci si concentra sull'*agente intelligente* come entità autonoma e fioriscono gli studi sui software intelligenti e sugli agenti intelligenti *embodied* in un sistema fisico². Incoraggiati dai risultati ottenuti con i computer di nuova generazione, negli scienziati si rinnova la speranza di costruire robot intelligenti, di inserire cioè una AI in un corpo sintetico per imitare l'essere umano. I progressi nella sensoristica consentono la creazione di sistemi di visione e di percezione molto avanzati; vengono realizzati sensori tattili e sistemi uditivi integrati in robot con capacità biomeccaniche sempre migliori. Gli algoritmi di controllo evolvono rapidamente. L'elettronica e i computer consentono di immagazzinare e utilizzare gli stimoli che provengono dai sensori per consentire al robot di muoversi in autonomia e, grazie alle prime AI *embodied*, persino di iniziare a prendere delle decisioni. Molti di questi sistemi si rivelano utili per risolvere problemi importanti per l'uomo e per la società³, ma fanno anche tornare in auge un'antica questione: possono gli esseri umani diventare obsoleti? Minsky aveva scritto: «I robot ereditano la terra? Sì, ma quei robot saranno i nostri figli!»⁴.

Per decenni corpo e mente erano stati studiati e imitati separatamente, dando origine a macchine molto potenti e precise, o a «cervelli elettronici» dalle straordinarie capacità computazionali. Con l'avvento delle macchine intelligenti (A/IS) si iniziano a integrare sistemi mecatronici con capacità sensoriali e prestazioni biomeccaniche simili a quelle del corpo umano, e sistemi di calcolo con potenze computazionali paragonabili o superiori a quelle del cervello umano. L'integrazione di questi due mondi tecnologici in macchine intelligenti costituite da robot dotati di AI costituisce un passo importantissimo nella storia della scienza e della tecnologia, e l'inizio di quella che si prospetta come una rivoluzione non solo tecnica ma anche antropologica, legale ed etica. Tuttavia, un'analisi oggettiva delle potenzialità e dei limiti degli A/IS mostra un ostacolo fondamentale che esse devono superare: la difficilissima riproduzione del nesso inscindibile fra corpo e mente, tipico dell'uomo e degli esseri viventi più evoluti.

1. MINSKY – PAPERT 1969.

2. Nel 1997 il computer Deep Blue dell'Ibm, capace di elaborare undici miliardi di operazioni al secondo, compie un'impresa storica battendo agli scacchi il campione mondiale Garry Kasparov. Allo stesso tempo risulta evidente che Deep Blue non può pensare e questa triste constatazione attenua l'entusiasmo per la ricerca sulla AI.

3. MCCORDUCK 2004; RUSSELL – NORVIG 2016; HUSBANDS 2014.

4. MINSKY – PAPERT 1969.

La distinzione fra corpo e mente può, in prima istanza, valere per la coppia robot-AI, ma certamente non è applicabile agli esseri viventi e, in particolare, all'uomo: l'orchestrazione corpo-mente dell'uomo e quella corpo-AI dell'umanoide sono completamente diverse. Nell'uomo, corpo e cervello sono profondamente interconnessi e sinergici; nessuno dei due ha un ruolo dominante: ciò che conta è l'armonica e perfetta interazione delle loro funzioni. L'apparato muscoloscheletrico umano si è evoluto di pari passo al sistema cognitivo, con un processo di adattamento reciproco, mediato dalla biochimica della vita (ormoni, metabolismo ecc). Lo stato emotivo condiziona la risposta fisica del corpo: la rabbia moltiplica le nostre forze, la paura ci fa vigili e concentrati e la tenerezza ci rende particolarmente delicati. Le fibre responsive che costituiscono i muscoli del corpo umano si contraggono e si rilassano grazie a stimoli nervosi consapevoli e inconsapevoli, correlati a stati mentali e a decisioni relative al nostro movimento o alle nostre necessità. Nel nostro corpo esistono sinergie sviluppatesi in miliardi di anni, che sfruttano meccanismi biologici, attualmente non riproducibili nei robot.

Un lungo lavoro d'integrazione è stato fatto, ma ancora c'è tanta strada da percorrere per rendere sinergici e compatibili l'attuatore motorizzato, che realizza il movimento, e il computer, che tale movimento deve comandare. Nella macchina intelligente una scheda elettronica computa algoritmi complessi che generano segnali digitali; questi, a loro volta, comandano interruttori e amplificatori che immettono correnti elettriche nei motori del robot, mediante un processo tuttora dispendioso e rudimentale rispetto a quello biologico. Nell'essere umano, al posto della corrente elettrica, flusso di elettroni, ci sono impulsi nervosi: pacchetti di ioni che si muovono in acqua, l'elemento che costituisce circa il 60% del corpo in un adulto. Il robot segue le leggi dell'elettricità, il corpo umano quelle della biochimica.

Quello descritto è, seppur per sommi capi, il processo tecnologico che ha dato origine alle macchine autonome e «pensanti». Per approfondire l'analisi è necessario illustrare alcuni dati notevoli per confrontarne le caratteristiche e performance con quelle dell'uomo. Innanzitutto, dobbiamo fare un raffronto che potremmo definire ecologico-costitutivo: l'uomo, come tutte le entità organiche, biologiche, naturali, è composto per il 99% da sei atomi: ossigeno (65%), carbonio (18%), idrogeno (10%), azoto (3%), calcio (2%) e fosforo (1%). Esso è progettato per crescere e, al termine della sua vita, dissociarsi in questi sei atomi. Qualunque macchina artificiale, invece, richiede dai 30 ai 50 atomi, è progettata per essere assemblata nel minor tempo possibile, e a fine ciclo vita qualcuno la dovrà smontare per recuperare i materiali. Per costruire un'automobile ci vogliono quattro ore, per disassemblarla 40, e si deteriorano molte parti.



Dal punto di vista del rapporto mente-corpo, l'uomo ha un sistema che si è ottimizzato in tre miliardi di anni di evoluzione: un lunghissimo lasso di tempo nel quale egli ha sviluppato straordinarie capacità di adattamento e di apprendimento. È ancora enorme il gap che la tecnologia deve superare per competere con i risultati dell'evoluzione umana. Il sistema di stabilità e di equilibrio dinamico dell'uomo è solo lontanamente imitabile dalla tecnologia: si può dare equilibrio a un robot utilizzando dei giroscopi (come quelli dei telefoni cellulari o degli aeroplani), ma i risultati non sono minimamente comparabili alle prestazioni del sistema vestibolare di un atleta o di un'acrobata. Se ci soffermiamo a considerare il nesso corporeo, scopriamo che esso è difficilmente eguagliabile dalla realizzazione di un computer connesso ad attuatori di movimento e a sensori. Grazie alla lunga e complessa evoluzione, oggi il cervello umano funziona in modo combinato e contemporaneo col corpo: lo stesso gruppo di neuroni che controlla la vista supervisiona anche l'attività di manipolazione; il gruppo che controlla la lingua supervisiona l'attività di comprensione del linguaggio, e così via. Attualmente è impossibile trasferire nelle macchine le sinergie mente-attuazione tipiche dell'uomo, perché l'intelligenza elettronica e il corpo mecatronico funzionano con meccanismi diversi da quelli biologici. Inoltre, è difficile pensare a macchine autonome intelligenti (A/IS) in grado di comunicare ad alto livello con gli uomini usando la semantica del corpo, interpretando correttamente i semplici gesti umani e capendone le intenzioni. Il linguaggio non verbale è parte importante della comunicazione umana: un ammiccamento, un volto contrariato possono dire più di mille parole; sono una modalità comunicativa estremamente intuitiva e veloce che usiamo continuamente, ma molto difficile da insegnare alle macchine.

Dobbiamo infine evidenziare un aspetto centrale nel paragone fra uomo e A/IS: questi ultimi hanno bisogno di una quantità di energia enormemente più grande (anche milioni di volte) per processare le istruzioni (Mips o le operazioni Peta Flop equivalenti) necessarie al funzionamento di un'entità biologica complessa come quella umana. Come evidenziato in precedenza, la capacità computazionale di alcune macchine può anche arrivare a centinaia di PetaFlop, tuttavia nella pratica hanno bisogno di potenze elettriche enormi (decine di milioni di watt), sono grandi come una stanza, e usano enormi centrali di raffreddamento e una centrale elettrica indipendente. Il cervello umano è una 'palla' che pesa meno di due chili, alimentata dal metabolismo degli zuccheri con poche decine di watt! Dunque, con gli standard tecnologici attuali è impossibile ipotizzare un sistema se-movente capace di pensare come l'uomo, con le stesse capacità mentali e biomeccaniche. Le Big Data Companies, che gestiscono i grandi computer



e i grandi apparati di storage, potrebbero offrire una parziale soluzione al dilemma apparentemente insolubile fra 'capacità intellettiva' e dispendio di energia da parte della macchina. I robot probabilmente resteranno 'stupidi' e con una capacità computazionale individuale limitata (intorno a un miliardo di operazioni al secondo, paragonabile a quelle di un buon computer che consuma qualche centinaio di watt). Questa ridotta 'intelligenza' potrà essere impiegata per muoversi bene, mentre tutta la parte 'cognitiva' dovrà essere gestita in modo differente. Si può ipotizzare la creazione di una mente unica a cui tutti i robot saranno collegati, una sorta di repository globale dell'intelligenza delle macchine, che utilizzerà il *cloud* per conservare tutte le informazioni e le 'cose imparate' dai robot; ciascuna macchina potrà fare l'*upload* delle proprie esperienze e il *download* di quelle altrui.

Si tratta di uno scenario affascinante, in cui alla memoria e all'intelligenza individuale degli esseri umani si contrappone una memoria e un'intelligenza unica e condivisa per i robot. Queste macchine potrebbero comportarsi come uno sciame, dotato di un'intelligenza unica cui contribuiscono tutti i suoi individui: una specie che non ha equivalente nel mondo biologico, e con la quale forse un giorno dovremo imparare a convivere.

Difficile dire quanto sia realistico questo scenario. Le difficoltà tecnologiche da affrontare sono ancora enormi: le tecnologie wireless velocissime, necessarie a dialogare in tempo reale con il cloud, ove dovrebbe risiedere l'unico grande intelletto a cui attingono tutti i robot (ricorda l'intelletto agente di Avicenna), non sono disponibili ovunque e richiedono infrastrutture di rete molto complesse (per esempio una rete 5G capillare). È probabile che siano molto più entusiaste e interessate a queste tecnologie le aziende che non costruiscono robot rispetto a quelle che lo fanno: queste ultime sanno che i robot da soli saranno sempre inferiori all'uomo, mentre le prime ritengono che una AI globale unica possa essere molto performante.

Va sottolineato che, pur con tutte le limitazioni appena esposte, già oggi si solleva un potenziale problema regolatorio ed etico: il giorno in cui si avranno macchine autonome e intelligenti (A/IS) sufficientemente sofisticate e computer sufficientemente potenti, con quali regole verrà gestita l'intelligenza globale del pianeta, il *global repository of intelligence* che diventerà il cloud? E chi lo gestirà? Gli stati oppure le grandi aziende? E qualora si riuscisse ad avere A/IS con una intelligenza elevatissima non condivisa nel cloud, ma residente individualmente in ciascuna macchina (scenario davvero improbabile al momento), come verranno trattate queste entità, a tutti gli effetti non biologiche ma capaci d'intendere e di volere?

Le macchine intelligenti potranno diventare capaci d'intendere e di volere ma, di sicuro, non divideranno la nostra stessa biologia. Sarà più semplice affrontare il problema pensando di avere a che fare con una razza aliena?



O dovremo cercare di adattare le regole che abbiamo sviluppato per noi stessi? Si tratta di questioni aperte a cui dobbiamo pensare. Senza catastrofismi o eccessivi ottimismo, ma intersecando la storia, la filosofia e le scienze umane con i nuovi orizzonti della tecnologia. Benché il giorno in cui potranno esserci robot intelligenti come l'uomo sembri lontano (e forse mai arriverà, a mio parere, almeno finché useremo il silicio), gli interrogativi di tipo etico, sociale, umano sollevati da una loro presenza ampia e diffusa nella società meritano attenzione.

Vorrei fornire alcune risposte o quantomeno indicare alcune linee guida per riflettere sull'argomento. Si tratta di considerazioni rivolte in particolare a coloro che sono coinvolti nella progettazione, nella realizzazione e nell'uso di questi nuovi rivoluzionari prodotti tecnologici. Avere macchine autonome e intelligenti (A/IS) in grado di agire sempre a favore dell'uomo e della collettività è un obiettivo che coinvolge esperti di molte e diverse discipline: ingegneri elettronici e meccanici, informatici, psicologi, neurologi, scienziati cognitivi, esperti di AI, logici, matematici, filosofi, giuristi, economisti, designer e artisti. Nel momento in cui un robot sviluppa la capacità di decidere e di operare in modo autonomo, di imparare, di acquisire esperienza con processi decisionali di natura algoritmica (benché in assenza di emotività e spiritualità), le problematiche etiche sono del tutto nuove e rilevanti. Una macchina cognitiva in grado d'imparare pone il problema di come educarla: quali strategie 'educative' mettere in atto; come promuoverla e come punirla in caso di violazione delle regole. La domanda fondamentale diventa: «Come si punisce un robot che sbaglia?»⁵. Gli uomini violano le regole per necessità, per errore, per vendetta o per cattiveria. Nella maggior parte dei casi la motivazione che conduce alla violazione è da ricondursi ad alterazioni psicologiche ed essenziali o a condizioni di particolare necessità o sofferenza. Il percorso di rieducazione passa attraverso una punizione che, normalmente, riguarda la riduzione di una libertà o si esprime in una sanzione da scontare. Più in generale, qualsiasi percorso educativo umano è fondato sul bilancio premio-punizione. La punizione si basa sul fatto che qualsiasi essere umano, e persino gli animali più intelligenti, temono di essere privati di qualcosa a cui tengono: sia essa la libertà, nel caso di una pena detentiva, o un giocattolo nel caso di un bimbo che abbia commesso una marachella. La paura di perdere qualcosa a cui si tiene è parte della psicologia umana, ma è anche conseguenza di quel principio di autoconservazione che anima tutti gli esseri viventi. Il timore di essere puniti ci fa comportare meglio per non subire peggioramenti della qualità della vita. Nella macchina, la violazione di una regola potrebbe scaturire semplicemente dalla valutazione che sia necessaria per ridurre alcuni effetti collaterali negativi di una determinata azione. La trasgressione sarebbe solo il frutto di un algoritmo che, minimizzando rischi e danni collaterali, mira al raggiungimento di uno scopo con delle condizioni al contorno.

5. WALLACH – ALLEN 2010.

Il caso più classico è quello di un'automobile a guida autonoma che, in caso di guasto ai freni, deve scegliere se sacrificare il passeggero fermandosi su un muro o sacrificare i pedoni proseguendo dritta nella sua direzione. A parità di violazione la macchina non potrà che scegliere la soluzione con la probabilità di danno minore, qualunque essa sia. Macchine diverse nella stessa situazione e con le stesse condizioni al contorno raggiungeranno tutte la stessa conclusione. Gli uomini, invece, a parità di situazione e di condizioni, non farebbero tutti la stessa cosa. Entrerebbero nella valutazione imponderabili elementi soggettivi che, inevitabilmente, porterebbero a decisioni individualmente diverse (ad esempio, se fra i pedoni ci fosse un parente la decisione finale potrebbe essere diversa, indipendentemente dal calcolo assoluto del rischio). La diversità dei comportamenti umani deriva dalla natura non algoritmica della nostra intelligenza, dalla componente sempre presente di irrazionalità, emotività, imponderabilità intrinseca ai nostri meccanismi logici. L'irrazionalità o non razionalità umana, frutto dell'ingrediente ormonale della nostra specie, genera creatività, fantasia, sentimento che, a loro volta, inducono variabilità e imprevedibilità nel nostro comportamento. Punire la macchina che ha violato una regola sarebbe molto difficile. Non le si potrebbe sottrarre cibo e libertà, in quanto in entrambi i casi essi non sono fondamentali per la sua specie. Né le si potrebbe privare della vita, poiché staccandole le batterie non faremmo nulla di diverso da quello che facciamo spegnendo un computer. La macchina potrà essere capace d'intendere e di decidere, ma ciò non basta ad avere una coscienza di sé e un conseguente istinto di conservazione, sopravvivenza e preservazione della specie. Qualunque sia il codice che si svilupperà per regolare la coesistenza fra gli umani e le macchine intelligenti, esso dovrà tenere conto del fatto che la macchina segue leggi diverse da quelle della biologia. Tutto ciò dipenderà da quanto rapidamente si evolverà l'identità e la personalità delle macchine autonome e intelligenti (A/IS). Nel caso, pur remoto, che esse divengano realmente una specie a sé stante, richiederanno un inquadramento giuridico ed etico specifico, che le inserisca all'interno della società e stabilisca per loro diritti e doveri. Non necessariamente un rapporto subalterno rispetto all'uomo, come l'animale sotto la responsabilità del padrone o l'artefatto sotto la responsabilità del costruttore, ma qualcosa di nuovo e profondamente diverso da quanto elaborato sino a oggi 

BIBLIOGRAFIA

- P. HUSBANDS, *Robotics*, in K. FRANKISH – W.M. RAMSEY (eds.), *The Cambridge Handbook of Artificial Intelligence*, Cambridge University Press, Cambridge 2014.
 P. MCCORDUCK, *Machines Who Think*, A.K. Peters Ltd., Natick, Massachusetts 2004.
 M. MINSKY – S. PAPERT, *Perceptrons: An Introduction to Computational Geometry*, MIT Press, Cambridge, Mass., 1969.
 S. RUSSELL – S. NORVIG, *Artificial Intelligence: a modern approach*, Pearson Education Ltd., New York 2016.
 W. WALLACH – C. ALLEN, *Moral Machines. Teaching Robots Right from Wrong*, Oxford University Press 2010.

* Approfondimenti sul tema nel saggio di D. ANDRESCIANI – R. CINGOLANI, *Robots and intelligent-autonomous systems: technology, social impact and open issues*, dal quale è stato rielaborato questo articolo, presentato al convegno: *Roboetica, persone, macchine e salute*, Pontificia Accademia per la Vita (Roma, 25 febbraio 2019).