



# L'ETICA AL TEMPO DEI ROBOT

Il robot, unione di mente sintetica e di corpo sintetico, rappresenta l'ultima versione del nostro tentativo plurisecolare di costruire l'uomo artificiale. La somiglianza sempre più spinta tra robot e uomo, che si estende alle capacità cognitive, all'autonomia e in prospettiva anche alle emozioni e forse alla coscienza, pone interrogativi inquietanti. La crescente diffusione dei robot in tutti i settori della società ci obbliga a considerare il rapporto di convivenza uomo-macchina in termini inediti, che coinvolgono in primo luogo l'etica. Affrontare questi problemi è importante e urgente.

## 1. INTRODUZIONE

**F**orse, in realtà, stiamo assistendo a una graduale fusione della natura generale delle attività e delle funzioni umane con le attività e le funzioni di ciò che noi umani abbiamo costruito e di cui ci siamo circondati.

Philip Dick, *Mutazioni*.

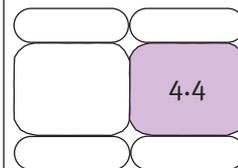
Mentre l'evoluzione biologica ha dotato gli organismi viventi prima di un corpo e poi di un cervello, avente funzioni di controllo centrale e dotato in certi casi di proprietà cognitive superiori, non strettamente necessarie alla regolazione del corpo, l'intelligenza artificiale funzionalistica ha invece cercato di costruire una mente senza corpo, cioè un'intelligenza che imitasse le funzioni simboliche e astratte del cervello biologico evitando ogni interazione con un ambiente considerato fonte di disturbo. Tuttavia, le difficoltà di estendere questa forma d'intelligenza artificiale al di fuori dei domini simbolico-formali, hanno fatto ritenere che soltanto accoppiando la mente artificiale all'ambiente, attraverso

un corpo artificiale dotato di sensi e di organi attuatori, si potesse ottenere un'intelligenza flessibile e ad ampio spettro come è quella biologica.

Il recupero della dimensione corporea e sensoriale ha portato ai robot e ha aperto una serie di interrogativi che vanno dagli aspetti tecnici della loro costruzione fino a sottili questioni di natura etica. Infatti, il robot è un artefatto capace di apprendere e dotato di una certa autonomia di decisione e comportamento e queste caratteristiche, in una prospettiva di stretta convivenza uomo-robot, non possono non sollevare certe domande come: fino a che punto siamo disposti a convivere coi robot, ad affidarci a loro nella vita quotidiana, nell'accudimento e nelle cure?

Se i robot dovessero un giorno diventare intelligenti e sensibili (quasi) quanto gli umani, potremmo continuare a considerarli macchine, come le lavatrici o le automobili? O dovremmo adottare atteggiamenti empatici e comprensivi come nei confronti degli animali domestici? Dovremmo arrivare a conferire loro dignità etica?

Giuseppe O. Longo



E viceversa: quali comportamenti dei robot dovremmo tollerare, incoraggiare o vietare? E di chi sarebbero le responsabilità di un loro eventuale comportamento dannoso?

L'ultima domanda è importante perché rivela il conflitto tra la natura artificiale dei robot, che dovrebbe renderli obbedienti alla nostra programmazione e la loro parziale autonomia (se un robot non è autonomo non è un robot) che, in linea di principio, potrebbe indurli a decisioni nocive nei nostri confronti. Il conflitto diventa drammatico nel caso dei "robot soldati". Erano problemi di questo genere che aveva in mente Asimov quando postulò le "Leggi della robotica", che vietano ai robot di compiere azioni dannose per gli esseri umani e che costituiscono il primo embrione di un'etica dei robot o, con un espressivo neologismo, di una "roboetica".

In questo ambito le previsioni si mescolano facilmente con la fantascienza e accanto alle speculazioni ci sono le realtà: in Giappone (il Paese di gran lunga più avanzato nella costruzione e nell'impiego dei robot) si tocca con mano quanto possa diventare intenso il rapporto uomo-macchina quando il robot sia un (o una) "badante" con sembianze umane oppure quando abbia più o meno le fattezze e il comportamento di un animale domestico (si pensi ad Aibo, il robot cane della Sony, ormai fuori produzione, che per anni ha svolto la funzione di "animale" da compagnia, Figura 1). La proiezione affettiva è tanto forte da suscitare problemi psicolo-

gici e, ancora una volta, etici. E poi, in generale, la marcia sempre più convulsa di una tecnologia invasiva e onnipresente non può non avere effetti profondi sull'immagine che abbiamo di noi stessi e sul nostro stesso essere "umani": specchiandoci in quello straniante *alter ego* che sta diventando il robot, quale immagine ce ne ritorna? Riusciremo, per differenza o per similarità, a capire qualcosa di più di noi stessi? Che questi problemi siano importanti e urgenti, è confermato dall'istituzione di un Comitato tecnico per la roboetica in senso alla *Robotics and Automation Society* dell'*IEEE*.

Nei paragrafi che seguono, dopo un breve inquadramento storico che descrive in particolare il passaggio dall'intelligenza artificiale funzionalistica alla robotica, sottolineando l'importanza del corpo sotto il profilo cognitivo e attivo, si considerano i problemi etici sollevati dalla presenza sempre più diffusa dei robot. Tali problemi sono acuiti dalla somiglianza crescente che presentano con gli umani, oggi sul piano cognitivo e attivo e, domani, forse, anche sul piano emotivo e della coscienza.

## 2. UNA STORIA MILLENARIA

L'impresa della robotica si colloca nel solco di un tentativo millenario, quello di imitare l'atto divino della creazione. Più o meno dichiarata, questa ambizione risale all'antichità biblica e classica, e la leggenda del Golem ne è forse l'esempio mitologico e letterario più noto (Figura 2). In questa impresa si intrecciano la vertigine del creatore e il timore per la creatura, che talora minaccia di ribellarsi e distruggere l'inesperto demiurgo. Anche nel caso del mostro di Frankenstein (Figura 3) la creatura trascende il progetto e si rivolta, suscitando negli uomini angoscia e terrore. Talvolta gli umani subiscono invece il fascino degli esseri artificiali: nei racconti di Hoffmann gli uomini si innamorano perdutamente di bambole meccaniche, imitazioni perfette della donna, in cui la differenza tra il modello e la sua riproduzione si attenua fino a scomparire, inducendo in inganno anche l'osservatore più attento. Invece, per l'imperizia del costruttore, il mostro di Frankenstein è segnato da una diversità



**FIGURA 1**

Il cane robot Aibo, di cui la Sony ha prodotto 150.000 esemplari tra il 1999 e il 2006

che suscita orrore perché è interpretata come segno di malvagità.

Questi temi - orgoglio e timore, fascino e orrore - sembrano appartenere a un passato ormai lontano, eppure a ben guardare sono ancora presenti non solo nelle opere di fantascienza, ma anche nell'immaginario collettivo e nel nostro atteggiamento nei confronti delle tecnologie di punta, in particolare delle "tecnologie della mente" come i computer, l'intelligenza artificiale e specialmente i robot. Ciò sembra confermare il profondo substrato mitopoietico ed emotivo che ha sempre accompagnato l'attività tecnologica e la nostra interazione con la macchina.

Accanto ai miti e ai racconti, l'ambizione di costruire l'uomo artificiale produsse nei secoli una fioritura di opere artigianali, gli *automi*, artefatti spesso zoomorfi o antropomorfi, che, mossi dalla forza idraulica, dalla gravità o da un meccanismo nascosto al loro interno, sembrano comportarsi come esseri viventi. Col tempo l'elemento meraviglioso e ludico fu sostituito dalle finalità pratiche: non si trattò più di costruire macchine che tentassero di compiere le mille diverse azioni di *un unico* uomo, ma al contrario di ottenere una macchina che compisse *un'unica* azione, però sostituendo *mille* uomini. È il passaggio dall'androide elegante e variopinto alla nera e possente macchina a vapore.

Gli automi, raffinati e suggestivi prodotti dell'ingegno umano, oggi non si costruiscono più e sono rimpiazzati dovunque, se non nei musei e nei teatri della nostalgia, dai robot, manufatti in cui la tecnologia punta sempre più all'efficienza e sempre meno all'imitazione puntuale della natura. Eppure gli automi, specie quelli antropomorfi, cioè gli androidi e le andreidi, continuano a popolare di inquiete proiezioni e torbidi sogni la dimensione immaginaria del nostro tempo e da qui, soffiati di suggestioni mitologiche, travalicano nelle creazioni artistiche (letterarie e cinematografiche) e nelle attuazioni tecniche (Figura 4). Tanto che anche la robotica si confronta con la costruzione di macchine antropomorfe, gli *umanoidi*, residuo di una storia affascinante e tenebrosa di meccanica onirica, dove magia e occultismo s'intrecciano con la genialità inventiva, in un turbinio di personaggi eterogenei, in-



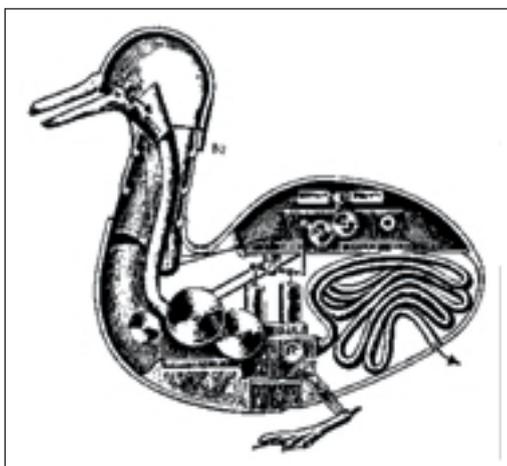
**FIGURA 2**

*Il Golem, costruito dall'uomo per farsi aiutare, a volte si ribella al suo creatore*



**FIGURA 3**

*Il mostro di Frankenstein nella celebre maschera cinematografica di Boris Karloff*



**FIGURA 4**

*L'anatra digerente, uno dei celebri automi di Jacques de Vaucanson (1709-1782)*

ventori, maghi, affaristi, ciurmadori, studiosi, prestidigitatori e creduloni.

### 3. DAL GOLEM AL FUNZIONALISMO

Per quanto stupefacenti, i prodotti artigianali del passato restavano comunque lontanissimi dal modello, uomo o animale, cui li avvicinava soltanto la forma esteriore ma non certo una puntuale somiglianza strutturale e funzionale. Le cose cambiarono radicalmente grazie alle ricerche stimulate dalla seconda guerra mondiale nel campo dei calcolatori e delle telecomunicazioni. Ben presto si capì che il calcolatore, lungi dall'essere una semplice macchina per far di conto, possedeva enormi capacità simboliche, tanto che nel 1956 nacque ufficialmente una nuova disciplina, cui fu dato il nome, un po' infelice per la verità, e fonte di equivoci durevoli, di *intelligenza artificiale* (IA) e il calcolatore divenne il modello di elezione della mente umana.

In fondo si trattava ancora della vecchia ambizione di imitare l'atto divino della creazione, ma non più con l'ingenuo e inarrivabile intento di costruire una creatura simile all'uomo nel suo complesso, magari con qualche approssimazione, bensì di riprodurre o simulare con estrema precisione una sola parte dell'uomo: la sua mente. Il sogno cominciava dunque a diventare realtà, sia pure limitatamente a un aspetto. Ma si trattava dell'aspetto più importante, caratteristico e fondamentale dell'uomo: l'intelligenza. Infatti, a quei tempi c'era (e c'è tuttora) una forte tendenza a identificare l'intelligenza con i suoi aspetti razionali, anzi simbolici e algoritmici, e questa identificazione, cui aveva contribuito potentemente il calcolatore, aveva a sua volta rafforzato la convinzione che l'informatica fosse la tecnologia giusta per costruire, dopo tante ingenuità, modelli della mente che fossero corretti e collaudabili.

Giungeva dunque a compimento un lunghissimo percorso, che dalla figura leggendaria del Golem portava, attraverso i mirabili e delicati automi, fino alla macchina ideale di Turing e ai prototipi concettuali di Von Neumann, capaci di riprodurre le funzioni nobili della mente. L'idea secondo cui, tutte le "funzioni nobili" della mente rientrassero nelle

possibilità di replicazione della macchina, per un verso restava avvolta nelle ambiguità definitorie e per un altro diveniva oggetto di una congettura, la tesi di Church, secondo la quale tutta l'attività mentale dell'uomo è di tipo algoritmico, dunque è riproducibile con una macchina discreta. Questa congettura, non dimostrata e non dimostrabile, fu accettata da molti e pose le premesse teoriche e la giustificazione filosofica della versione *forte* dell'IA. Secondo questa versione, è possibile trasferire, senza perdite e senza distorsioni, da una struttura (cervello) a un'altra (computer) la funzione (cioè i programmi e gli algoritmi), che è la vera essenza dell'intelligenza. Si parla perciò di funzionalismo.

### 4. L'ALTRA METÀ DEL ROBOT: IL CORPO

Dopo i primi lusinghieri successi, anche i sostenitori più ferventi del funzionalismo dovettero riconoscerne i limiti, che derivano dalla natura disincarnata della mente artificiale, cioè dall'assenza di un *corpo* che comunichi con l'ambiente. Se l'intento era quello di simulare l'intelligenza umana, il *riduzionismo mentalista* dell'IA funzionalista ne trascurava un elemento essenziale.

L'intelligenza umana (e animale) si costituisce e si manifesta attraverso il corpo. L'intelligenza è un insieme di caratteristiche e attività fortemente sistemiche, oltre che fortemente dia-croniche, cioè evolutive. In particolare, l'intelligenza nasce, si sviluppa e si manifesta attraverso la *comunicazione*, cioè lo scambio di messaggi di vari tipi, entro vari contesti, in vari codici e a vari livelli. Poiché la nostra "interfaccia" con il resto del mondo è costituita dal corpo e dagli strumenti tecnologici che abbiamo via via creato e perfezionato e che del corpo sono un potenziamento e un'estensione, è chiaro che proprio al corpo spetta il compito determinante di consentire la comunicazione e di filtrarla, sia in ingresso sia in uscita.

Riconosciuto il limite essenziale del funzionalismo e proseguendo sulla strada dell'imitazione della natura, si trattava di dotare il cervello artificiale di un corpo artificiale: questa strada portò alla robotica. Alla base di questa svolta c'è il riconoscimento della funzione conoscitiva del corpo.

Il sistema o macchinario conoscitivo individuale ha due modalità essenziali di funzionamento. La prima, più arcaica sotto il profilo sia filogenetico (della specie) sia ontogenetico (dell'individuo), è la conoscenza tacita, globale e immediata attuata dal corpo, nella sua struttura e nelle sue funzioni biologiche, e guidata dal sistema affettivo ed emotivo. La seconda, più recente sotto il profilo evolutivo e posteriore nello sviluppo dell'individuo, è la conoscenza esplicita, attuata nelle forme verbali e della razionalità. La prima è una conoscenza che si attua nel corpo e tramite il corpo, la seconda si attua nella mente o tramite la mente.

Orbene, la storia della cultura occidentale, in particolare della scienza, è in fondo un lungo tentativo di trasferire le conoscenze dalla prima alla seconda modalità, cioè dalla conoscenza biologica incarnata nel corpo (corpo che a sua volta è immerso nell'ambiente) a una razionalità disincarnata. In altre parole si vorrebbe tradurre nello scarnificato linguaggio astratto della mente (in particolare nel simbolismo della matematica) le rigogliose strutture del corpo e in genere della realtà; di rendere cioè esplicito, consapevole e leggibile ciò che è implicito, inconsapevole e oscuro. Questo tentativo è culminato nell'impostazione funzionalista o fisico-simbolica dell'IA. Ma fino a che punto è possibile questo trasferimento? All'inizio si riteneva che tutte le conoscenze fossero trasferibili, ma dopo i primi entusiasmi sono venute le delusioni e oggi ci si rende conto che, per replicare compiutamente l'intelligenza umana (ammesso che sia questo lo scopo dell'IA), anche le macchine intelligenti non possono fare a meno dell'equivalente di un corpo con tutta la sua attività cognitiva profonda e in parte forse non algoritmica: l'intelligenza disincarnata è troppo fragile e limitata.

Insomma, il tentativo di tradurre in conoscenza alta, razionale ed esplicita la massa delle conoscenze materiali, corporee e implicite incappa nell'ostacolo tipico di ogni processo di traduzione, cioè l'*incompletezza*. Rimane sempre un residuo ostinato e ribelle, che non si può tradurre.

Abbiamo così giustificato in termini epistemologici il passaggio dall'IA funzionalistica all'IA incorporata nel robot. Questo passag-

gio si oppone in qualche misura a una lunga tradizione filosofica. Da Platone in poi la modalità di conoscenza razionale è stata considerata superiore a quella corporea e tutta la corrente filosofica dominante si iscrive in questa ottica. Nel solco della filosofia razionalistica, anche l'IA funzionalistica considera la conoscenza astratta più nobile di quella legata al senso comune: l'intelligenza che dimostra un teorema sarebbe superiore a quella che riconosce una scena o che ci guida nelle azioni quotidiane. Ma questa lunga tradizione oggi viene messa in discussione. Addirittura si assiste a un capovolgimento: si riconosce che la maggior parte delle conoscenze, specie quelle vitali, sulle quali poggiano e dalle quali scaturiscono tutte le altre, sono espresse nella struttura stessa del corpo. A sua volta il corpo è immerso in un ambiente il quale, con le sue continue perturbazioni, non cessa di arricchire e aggiornare quelle conoscenze.

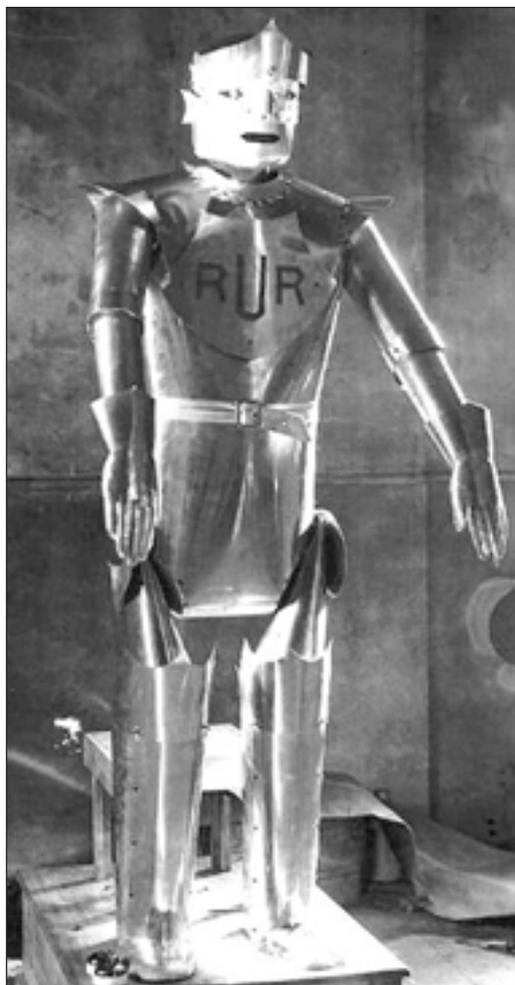
Il futuro della robotica più ambiziosa, quella che mira alla costruzione di macchine dotate di intelligenza, emozioni e forse coscienza, potrebbe dunque dipendere dalla comprensione del significato cognitivo delle azioni semplici, incarnate e contestualizzate che compiamo di continuo nella vita di tutti i giorni. Le descrizioni e gli strumenti usati finora in IA sono "alti e deboli": occorre integrarli con descrizioni e strumenti "bassi e forti", che riflettano e riproducano il nostro sfuggente "essererci nel mondo".

## 5. L'URGENZA DI UNA RIFLESSIONE

Per le considerazioni che intendo svolgere sui robot, è utile prendere spunto dalla letteratura fantascientifica, che costituisce un importante laboratorio di scenari suscettibili di trasformarsi in realtà, se non nei particolari, almeno nei tratti generali. Nel caso dei robot la letteratura e la filmografia sono ricchissime di spunti. La psicologia e la sociologia dei robot, degli androidi e dei ciborg (o *cyborg*, all'inglese) sono uno dei temi più interessanti della fantascienza moderna e si annunciano come uno dei settori più problematici di un futuro già a portata di mano nell'ambito della robotica. A cominciare dal dramma *R.U.R.* di

### R.U.R.

Il termine "robot" (dal vocabolo ceco "robota", ossia lavoro pesante, sfacchinata) fu introdotto nel 1920 dallo scrittore ceco Karel Čapek nel suo dramma "R.U.R." (*Rossum's Universal Robots*, rappresentato il 25 gennaio 1921 al Národní divadlo di Praga) per indicare una macchina antropomorfa progettata e costruita dall'ingegner Rossum (da un'altra radice slava che significa "intelligenza") per alleviare le fatiche degli umani. Nel dramma si ritrovano molti dei temi relativi al rapporto uomo-robot: la compassione di Helena, che li ritiene infelici e vorrebbe promuoverne il riscatto dotandoli di anima; il realismo di Domin che li considera semplici macchine, prive di ogni sensibilità e destinate a servire indefessamente gli umani; il cinismo di Gall, che le vorrebbe capaci di soffrire per aumentare il loro rendimento; la ripugnanza di Nana, che vede in loro l'opera del demonio; gli effetti perversi della loro laboriosità, che porta gli uomini ad affogare nell'ozio e le donne a non partorire più; il loro impiego militare contro gli operai in rivolta per aver perso il lavoro. Per opera degli scienziati, i robot progrediscono e diventano sempre più intelligenti, superando gli uomini. Quando se ne rendono conto, i robot di tutto il mondo si ribellano ed eliminano la razza umana per assumere il potere, però così facendo si condannano alla scomparsa perché senza gli uomini non sanno riprodursi. Ma due robot di tipo specialissimo, maschio e femmina, hanno ricevuto dai costruttori scomparsi la capacità di amare e di procreare e danno origine a una nuova stirpe.



**FIGURA 5**

Un robot del dramma R.U.R. (*Rossum's Universal Robots*) dello scrittore ceco Karel Čapek in un allestimento del 1930 circa

Karel Čapek (*riquadro* - Figura 5), passando per racconti di Isaac Asimov dedicati ai robot, fino a film come *2001: Odissea nello spazio*, *AI: Intelligenza Artificiale*, *Blade Runner* e molti altri ancora, scrittori e registi hanno indagato con slancio e inventiva il rapporto uo-

mo-macchina, indicandone i possibili sviluppi e i nodi prossimi venturi (Figura 6).

Questi scenari possiedono ovviamente una forte tinta fantastica, ma rispecchiano tendenze e problemi che negli ultimi anni, da una parte si sono sempre più avvicinati alla realtà della vita quotidiana e dall'altra, tendono a imprimare su di sé molte ricerche nell'ambito dell'IA, della robotica, della protetica, dell'ibridazione nanometrica. È alla luce di questi sviluppi e di queste tendenze che si devono considerare le prospettive aperte dalla robotica e dalla presenza dei robot tra noi. I robot lavorano in collaborazione con noi, una collaborazione che per il momento si configura come dipendenza, ma che in un qualsiasi momento del futuro potrebbe assumere carattere paritario per i continui progressi tecnici. La distinzione, oggi chiarissima, tra uomo e robot tende ad attenuarsi, l'antropologia tende a confondersi con la "robotologia".

A questo proposito si prospetta il vasto problema della sostituibilità del robot all'uomo, problema che fu già affrontato da Norbert Wiener quando si rese conto delle possibili implicazioni della tecnologia dell'informazione. Ma la riflessione filosofica risale addirittura a Kant, che pose un chiaro divieto all'uso strumentale dell'essere umano. Il problema della sostituibilità ha un aspetto tecnico (si tratta di valutare il rapporto mezzi-fini in un contesto specifico); un aspetto economico (rapporto costi-benefici, che nel caso dei robot di servizio, badanti, camerieri e così via, potrebbe comprendere una valutazione più soggettiva, legata alla cortesia, alla gradevo-

lezza, all'estetica); un aspetto legale (d'importanza cruciale è l'attribuzione della responsabilità di un danno provocato da un robot nell'interazione con una persona). E infine: esistono settori in cui l'integrazione o la sostituzione sia da escludere? Questa domanda apre la prospettiva etica nel senso kantiano, che ha a che fare con la dignità, con i fini, con l'etica e che a sua volta dovrebbe essere la base per le decisioni politiche e, in ultima istanza, anche tecniche.

Tra gli specialisti è diffusa l'opinione (spesso irriflessa) che l'introduzione nella nostra vita di macchine intelligenti (qualunque sia il significato di questo aggettivo) e la sostituzione di queste macchine al posto degli esseri umani portino benefici generalizzati alla società. Questa opinione dovrebbe confrontarsi con un fatto che è sotto gli occhi di tutti: sempre più le innovazioni assumono un carattere imperativo. Cioè si diffondono in base a una motivazione implicita e intrinseca, di carattere tecnico-economico e non perché una discussione aperta e democratica abbia stabilito che sono vantaggiose, magari dopo un periodo di assestamento in cui certi settori potrebbero risultare danneggiati.

Poiché la tecnologia avanza con velocità crescente, è opportuno dedicare attenzione a questi problemi, che sono complicati dal fatto che nel rapporto uomo-macchina è l'uomo che, per la sua flessibilità, di solito si adatta alla tecnologia e non il contrario. Questo adattamento comporta trasformazioni antropologiche che da alcuni, impropriamente, sono state assimilate a una "disumanizzazione". In realtà si tratta di modifiche di tipo evolutivo, e ve ne sono sempre state. Il vero problema è che la loro velocità e il loro susseguirsi rende spesso la trasformazione dolorosa. I problemi indotti dalla stretta interazione, o meglio dalla simbiosi, tra l'uomo e la macchina sono di natura non solo cognitiva, culturale o sociale, ma, anche e squisitamente, etica. È abbastanza singolare che il dibattito etico si accenda intorno alle innovazioni biologiche, genomiche e procreative, mentre sul fronte della tecnologia basata sull'IA, come la robotica, si osserva una tacita accettazione del fatto compiuto. Ma poiché, come cercherò di argomentare, gli effetti delle macchine intelli-



**FIGURA 6**

*Una scena del film AI: artificial intelligence di Steven Spielberg (2001)*

genti sono di vasta portata, essi esercitano una forte pressione sull'etica. Il problema etico, già di per sé arduo nel mondo di oggi, viene complicato da questi nuovi attori che sono i robot: è importante capire che cosa si fa, come lo si fa e perché lo si fa, tenendo conto, per quanto possibile, delle conseguenze delle scelte compiute oggi, conseguenze che potrebbero dimostrarsi irreversibili. È un richiamo alla nostra responsabilità, che a sua volta deriva da una consapevolezza che ormai si fa strada tra i ricercatori più avvertiti.

## 6. ETICA ED ESTETICA

Il tema della ribellione della creatura nei confronti del creatore è una costante dei rapporti uomo-tecnologia e ha molti precedenti nella tradizione. L'inquietudine derivante dalla possibile insubordinazione affiora anche oggi, forse perché i robot ci imitano nelle funzioni e nel comportamento e potrebbero diventare nostri concorrenti. La somiglianza delle forme acuisce l'inquietudine: un robot a forma di frigorifero non c'impresiona quanto un umanoide, anche se meno "intelligente" del primo. All'umanoide tendiamo ad attribuire caratteristiche umane (intelligenza, sentimenti...) che esiteremmo a concedere ai robot non antropomorfi. Le suggestioni derivanti dalla somiglianza esteriore di forma sono fortissime e

formano un cortocircuito destabilizzante quando si scontrano con la consapevolezza che ci si trova di fronte a una macchina (Figura 7). Ciò che si sa per via razionale rischia di essere spazzato via dalla proiezione emotiva: il robot viene umanizzato grazie a un meccanismo simile a quello che ci fa attribuire alle menti altrui, inaccessibili, le stesse proprietà della nostra mente, che ci è un po' più accessibile. È una sorta di animismo, un'estensione ai manufatti artificiali dell'antropomorfizzazione che esercitiamo da sempre nei confronti dell'alterità (per esempio divina o animale).

Ciò conferma quanto siamo sensibili all'aspetto esteriore delle creature che ci circondano: l'estetica è sempre stata una guida importante per le nostre azioni e per le nostre scelte (per esempio in campo sessuale e procreativo). Inoltre etica ed estetica sono legate a doppio filo: ciò che è bello ci appare anche buono e viceversa (l'endiadi greca *kalòs kài agathós*, bello e buono, la dice lunga). Etica ed estetica affondano le loro radici nella nostra storia evolutiva, anzi nella coevoluzione tra noi e l'ambiente. Propongo le seguenti definizioni naturalistiche, che si basano su una impostazione sistemica simile a quella di Gregory Bateson:

□ *l'estetica è la percezione soggettiva (ma condivisa) del nostro legame con l'ambiente, legame caratterizzato da una profonda ed equilibrata armonia dinamica;*



**FIGURA 7**  
*Valerie, una androide che la pubblicità dichiara utile per sbrigare le faccende domestiche*

□ *l'etica è la capacità, soggettiva e intersoggettiva, di concepire e compiere azioni capaci di mantenere sano ed equilibrato il legame con l'ambiente.*

Etica ed estetica sono due facce della stessa medaglia perché derivano dalla forte coimplicazione evolutiva tra specie e ambiente e sono entrambe "rispecchiamenti" in noi di questa coevoluzione. Se l'estetica è il sentimento (inter)soggettivo dell'immersione armonica nell'ambiente e l'etica è il sentimento (inter)soggettivo di rispetto per l'ambiente e di azione armonica con esso, allora l'etica ci consente di mantenere l'estetica e l'estetica ci serve da guida nell'operare etico.

Si noti che l'etica tradizionale è molto più ristretta di quella proposta qui, poiché si limita a considerare i rapporti tra esseri umani. In questo senso le religioni hanno costruito etiche codificate (o morali) basate sul rapporto interpersonale mediato o imposto dalla divinità. È anche interessante notare che l'attenzione quasi esclusiva della morale religiosa per l'uomo ha portato al concetto di persona, all'idea di dignità dell'uomo, e alla formulazione di vari codici o statuti dei diritti dell'umanità. Non voglio affatto sminuire la portata di queste conquiste, anzi forse se ci si limitasse all'impostazione naturalistico-evolutiva da me proposta non si riuscirebbe a fare il salto qualitativo compiuto dalla morale religiosa e incorporato nella legislazione di molti Paesi. È necessario tuttavia notare che l'attenzione per l'uomo è oggi affiancata da una crescente attenzione per alterità non umane, ad esempio per gli animali e per l'ambiente, attenzione che recupera in parte la radice naturalistico-evolutiva dell'etica. È in questo quadro allargato che si può immaginare di elaborare un'etica che comprenda anche i rapporti tra uomo e robot, oggi in via speculativa ma tra pochissimo anche in pratica. Infatti, l'equilibrio del sistema complessivo, di cui facciamo parte, e che sta alla base della definizione di etica e di estetica, è dinamico, non certo statico: perciò etica ed estetica sono *storiche ed evolutive* e dipendono anche dagli oggetti artificiali che l'uomo costruisce e che sempre più concorrono a formare l'ambiente in cui viviamo. Di questo ambiente cominciano a far parte anche i robot, quindi è inevitabile che essi influiscano sulle nostre percezioni



estetiche e sui nostri valori etici, dunque sul complesso dei nostri comportamenti.

## 7. LA ROBOETICA

*Che siamo fatti di carbonio o di silicio non ha importanza: ciascuno di noi deve essere trattato col giusto rispetto.*

Arthur C. Clarke, 2010

Esaminiamo ora il concetto di “roboetica”, cercando di esplicitarlo nei suoi significati possibili. Dalle considerazioni precedenti emerge una prima accezione, molto generale: “roboetica” è semplicemente “l’etica nell’epoca dei robot”, cioè l’insieme dei comportamenti dell’umanità quando anche i robot fanno parte dell’ambiente. “Roboetica” potrebbe anche significare l’insieme (molto più ristretto del precedente) di quei nostri comportamenti nei confronti dei robot che consentono di mantenere un giusto equilibrio dinamico tra noi e loro. Poiché i robot posseggono una certa autonomia e una certa capacità di apprendere dall’esperienza, “roboetica” può anche indicare l’insieme dei comportamenti utili, o almeno innocui, dei robot nei *nostri* confronti. Infine, ed è il significato più avveniristico, potrebbe significare il complesso dei comportamenti che i robot adottano tra loro e verso il loro ambiente, di cui fanno parte anche gli umani.

Riassumendo, la roboetica può significare:

- a.** *l’etica umani* → *ambiente (ambiente in cui ci sono altri umani e anche i robot);*
- b.** *l’etica umani* → *robot;*
- c.** *l’etica robot* → *umani;*
- d.** *l’etica robot* → *robot + ambiente (ambiente in cui ci sono anche gli umani).*

Mi rendo conto che si tratta di definizioni approssimative e discutibili, ma da qualche parte bisogna pur cominciare. La terza accezione (punto *c*) si deve conformare al precetto generale e tradizionale per cui le macchine non debbono danneggiarci (*primum non nocere*). È a questo proposito che Isaac Asimov (Figura 8) propose, in un racconto del 1942, le sue famose “Leggi della Robotica”, le quali, cablate in modo inestirpabile nel cervello positronico dei robot, dovrebbero tutelarci da comportamenti ostili e dannosi:

**1.** *un robot non può recar danno a un essere umano e non può permettere che, a causa di*

*un suo mancato intervento, un essere umano riceva danno;*

**2.** *un robot deve obbedire agli ordini impartiti dagli esseri umani, purché tali ordini non contravvengano alla Prima Legge;*

**3.** *un robot deve proteggere la propria esistenza, purché la sua autodifesa non contrasti con la Prima o con la Seconda Legge.*

Queste tre Leggi si presentano semplici, chiare e univoche: dovrebbero bastare per regolare perfettamente almeno il punto *c*. In realtà se le regole di Asimov fossero calate nel mondo reale non mancherebbero di suscitare problemi e ambiguità. Che cosa vuol dire danno? Chi ne è responsabile? E chi lo stabilisce, chi lo quantifica? Il concetto di danno sembra legato al concetto di male (non solo fisico) e sul problema del male si sono arrovellate generazioni di filosofi, teologi, letterati e artisti. Il cervello positronico, razionale e rigoroso, saprebbe impostare e risolvere le “equazioni del male” grazie a un’edizione aggiornata del *calculemus* leibniziano? C’è da dubitarne.

In effetti la nozione di danno che compare nelle Leggi, presenta molte ambiguità: se un umano sta recando danno a un altro essere umano (per esempio sta tentando di ucciderlo), come si deve comportare il robot? Se interviene reca danno all’assassino, ma il suo mancato intervento reca danno alla vittima.



**FIGURA 8**

*Lo scrittore Isaac Asimov, autore delle celebri Leggi della robotica*

Inoltre noi uomini siamo contraddittori: come si deve comportare un robot che riceva un ordine contraddittorio (dallo stesso uomo o da due uomini diversi) che sotto il profilo logico metta in crisi il suo sistema di valutazione? Di fronte a una contraddizione gli umani se la cavano quasi sempre con scelte che li fanno "uscire dal sistema" all'interno del quale si annida la contraddizione. Ma questa evasione (che corrisponde forse all'ampliamento dei sistemi formali entro i quali si riscontrano le limitazioni di tipo gödeliano) può avvenire grazie a una certa dose di irrazionalità o di follia creativa. Per consentire al robot di non paralizzarsi di fronte a una contraddizione si potrebbe forse immaginare di iniettarli una certa dose di follia, ma si può immaginare la difficoltà di un'impresa del genere.

Si può continuare a speculare: se si affidasse lo sviluppo della "specie" robot a un processo evolutivo analogo a quello biologico (o a quello bio-culturale), essi potrebbero compiere - in sostanza fuori del nostro controllo - progressi tali da consentir loro valutazioni etiche più raffinate e precise delle nostre. Potrebbero, prima o poi, cavarsela meglio di noi in tema di bene e di male (anche se il bene e il male sono sempre riferiti a un soggetto: bene per chi? Male per chi?) e potrebbero sviluppare una "teodicea" più rigorosa e soddisfacente della nostra, cioè potrebbero avvicinarsi alla soluzione di un problema teologico e metafisico che ci assilla da sempre: se il creatore del nostro mondo è bontà infinita, perché nel mondo c'è il male? Ma a quel punto dovrebbero ancora sottostare alla prima Legge? Oppure sarebbero loro a dettarci leggi nuove e ad assumere il bastone del comando, come solerti genitori nei confronti dei loro vivaci e stolti frugoletti? Del resto in *2001: Odissea nello spazio* il calcolatore Hal 9000 si comporta proprio così: prende il comando della nave e tenta di uccidere gli umani che intralciano il compimento della missione, invertendo l'ordine d'importanza delle Leggi, cioè subordinando la Prima e la Seconda alla Terza.

Asimov si era certo posto problemi di questo tipo, tanto che in seguito aggiunse la Legge Zero:

*o) un robot non può recar danno all'umanità e non può permettere che, a causa di un suo mancato intervento, l'umanità riceva danno.*

L'ultima Legge è interessante per il suo carattere "meta" e conferma che le prime tre non sono sufficienti a costituire un'etica di tipo c sicuro. Infatti se un folle minacciasse la distruzione in massa dell'umanità, la Legge Zero autorizzerebbe il robot a eliminarlo, contro la Prima Legge. Si apre qui il problema della valutazione quantitativa dei danni, ragionevole anche se molto discutibile secondo la morale tradizionale: l'uccisione di molti è (sarebbe) più grave dell'uccisione di uno.

Ma neppure con quest'aggiunta le leggi di Asimov riuscirebbero a proteggerci da comportamenti robotici dannosi, perché le conseguenze ultime di un'azione, pur rispettosa delle quattro leggi, potrebbero alla lunga essere nocive, e l'analisi di queste conseguenze di lunga portata sfiderebbe la più potente intelligenza (naturale o artificiale) immaginabile: troppe sono le ramificazioni e le interazioni con la mutevole complessità del reale. Del resto anche le azioni umane dettate dalle migliori intenzioni del mondo sfociano spesso in disastri. Inoltre ci si può chiedere dove ci si debba arrestare nella catena delle conseguenze di un'azione per valutare se l'azione sia stata buona o cattiva. Nella società umana solo alcune azioni "cattive" sono giudicate tali esplicitamente e sono sanzionate in un momento preciso grazie a un processo giudiziario che interrompe (o almeno vorrebbe interrompere) la catena delle causazioni: la maggior parte dei nostri atti non sono oggetto di giudizio formale a un istante dato e continuano a provocare conseguenze nel mondo ben al di là delle nostre intenzioni e per un tempo potenzialmente illimitato.

## 8. I ROBOT SOLDATO

*A quale crocevia l'evoluzione in noi umani ha imboccato la strada sbagliata, al punto che abbiamo associato il soddisfacimento del piacere alla spinta alla distruzione?*

Christa Wolf, *Guasto*

*Secondo me non ci siamo arrivati: è innato nella nostra specie. Il desiderio di distruzione è così radicato in noi che nessuno riesce ad estirparlo. Fa parte della costituzione di ognuno, giacché il fondo dell'essere stesso è*

*certamente demoniaco. Il saggio è un distruttore placato, in pensione. Gli altri sono distruttori in servizio.*

E.M. Cioran,  
*Dell'inconveniente di essere nati*

Un esempio già attuale di problema roboetico è rappresentato dall'uso in guerra dei robot soldato, cioè di robot costruiti, addestrati e impiegati in azioni belliche, con lo scopo precipuo di uccidere i nemici (Figura 9).

La Prima Legge impedirebbe ai robot di partecipare ad azioni belliche contro esseri umani, mentre oggi molte ricerche mirano proprio alla costruzione di robot soldato. Queste ricerche sembrano trovare qualche giustificazione, almeno in certi casi, nella Legge Zero, che autorizzerebbe a recare un certo danno (a uccidere *alcuni* umani) a chi vuole provocare danni ultimi e irreversibili (uccidere *tutti* gli umani). Come ho detto, si intravede qui una *scala quantitativa* dei danni, che relativizza il carattere in apparenza assoluto delle Leggi e conferma la difficoltà della loro applicazione. Le ricerche sui robot da guerra s'inseriscono nel quadro del combattimento a distanza, che aumenta l'efficienza e ottunde la pietà nei confronti del nemico. L'inserimento tra me e il nemico di un robot soldato aggiunge alla distanza fisica un distanza psicologica che colora la battaglia di indifferenza, di cinismo e di irresponsabilità. Quest'ultimo punto è forse il più importante: delegando al robot l'uccisione del nemico, l'uomo si scaricherebbe in buona parte della responsabilità del sangue versato. Ma fino a che punto la responsabilità di un'azione criminosa può ricadere sulla "macchina" robot, che almeno per il momento non ha statuto giuridico? Solo nell'ipotesi che il robot possedga una volontà autonoma e magari una coscienza riflessiva si può pensare a un'attribuzione di responsabilità. Altrimenti essa continua a ripartirsi tra progettisti, costruttori, militari e politici. È evidente che si tratta di un problema etico di tipo misto *a e c*. Infatti, la battaglia è un'impresa voluta da umani contro umani, ma è mediata e condotta da robot (semi)autonomi. Osserviamo di passaggio che un robot soldato, anche se votato ad azioni di morte può conservare un residuo di eticità: anche se svincolato dalla Prima Legge, il robot dovrebbe poter riconoscere un nemico che si arrende o



non è più in condizioni di combattere, in modo da farlo prigioniero invece di ucciderlo.

Col tempo gli umani hanno sviluppato codici di comportamento nei confronti dei nemici o dei prigionieri che aprono isole di misericordia nell'ambito della crudeltà bellica. Si apre qui l'analogo problema per i robot: come indurre nei robot comportamenti di compassione o in genere di etica bellica nei confronti degli umani? La domanda rivela il conflitto tra la loro natura meccanica, che dovrebbe renderli obbedienti alla nostra programmazione, e la loro (parziale) autonomia che, in linea di principio, potrebbe indurli a decisioni nocive nei confronti degli uomini oltre quelle codificate dalle convenzioni belliche (*riquadro*).

#### FIGURA 9

*Un robot soldato montato su cingoli*

#### I robot soldato

I tentativi di far condurre le operazioni militari alle macchine non è certo nuovo. Lo scopo è quello di infliggere perdite al nemico risparmiando i propri combattenti. Già nella seconda guerra mondiale i tedeschi usarono i Goliath, piccoli carri armati telecomandati, i missili Cruise non hanno pilota e si dirigono con buona precisione sul bersaglio. Ora gli Stati Uniti costruiscono robot con funzione di spionaggio e di combattimento, i cosiddetti SWORDS (spade, ovvero *Special Weapon Observation Reconnaissance Detection Systems*), dispositivi con mitragliatrice telecomandati fino a un chilometro di distanza. Gli SWORDS sono un primo passo, per quanto modesto, verso i *Future Combat Systems* (FCS), complessi di sorveglianza e attacco a distanza con missili e cannoni. I robot soldato si muovono su cingoli, ruote o gambe snodate e possono essere impiegati anche per il salvataggio di feriti e il recupero di materiale. Un altro settore in cui si prospetta l'impiego dei robot soldato è quello della lotta al terrorismo e della guerriglia urbana. La Francia si è impegnata nella costruzione di un robot antisommossa e a Singapore si stanno studiando robot soldato per combattere la criminalità urbana riducendo le perdite tra le forze dell'ordine.

## 9. SIMILI A NOI?

*Dottor Gall: i Robot quasi non avvertono i dolori fisici. Ciò non ha dato buoni risultati. Dobbiamo introdurre la sofferenza.*

*Helena: e sono più felici se sentono il dolore?*  
*Dottor Gall: al contrario; però sono tecnicamente più perfetti.*

Karel Čapek, *R.U.R.*

I problemi di tipo *d* sono certo quelli più avveniristici, e li possiamo tralasciare, mentre non sono così lontani nel futuro quelli di tipo *b*, che riguardano il nostro comportamento verso i robot. Negli ultimi tempi si è acuita in molti Paesi la sensibilità nei confronti degli animali superiori, come le scimmie e gli animali domestici, ma non solo. Ne sono prova la nascita di associazioni animaliste e di movimenti antivivisezione, la diffusione dell'alimentazione vegetariana e il crescente rifiuto di pellicce, avorio e altri "prodotti" animali. Questa maggior sensibilità è forse legata a un progressivo affrancamento degli animali dal ruolo di schiavi, di forza lavoro e di riserva di materiali utili cui sono stati a lungo relegati, ruoli che si sono trasferiti alle macchine e ai prodotti di sintesi. A riprova si rifletta che le bestie allevate a scopo alimentare non beneficiano ancora di questo incremento di compassione. Dell'affrancamento godono via via anche gli schiavi umani (spesso trattati come animali), non appena le loro funzioni si possono trasferire alle macchine. E qui entrano in scena i robot, che stanno diventando gli esecutori di molti dei lavori finora svolti dagli animali, dagli schiavi e dalle macchine tradizionali. Può accadere che la sensibilità diffusa nei confronti degli umani e degli animali si trasferisca prima o poi anche ai robot, oppure ai nostri occhi prevarranno sempre la loro natura di macchine e la loro funzione servile? Gli sforzi che facciamo per dotarli di intelligenza, autonomia, capacità di apprendere e tendenzialmente anche di sensibilità e coscienza, avranno come corollario una loro equiparazione a qualcosa di più nobile e vicino a noi? Ma c'è un'altra domanda, più inquietante: che diritto abbiamo di costruire macchine tanto intelligenti e sensibili da capire che non lo sono abbastanza? Perché suscitare dal nulla creature tanto simili a noi da essere capaci di soffrire? Il loro dolore, scaturito dalla coscienza di non essere del tutto assimila-

bili agli uomini, sarebbe un triste corollario della nostra abilità demiurgica: creando una schiatta di "macchine dolenti", ci assumeremmo una pesante responsabilità (*riquadro a p. 15*).

Le stesse domande si possono porre, e forse con fondamento ancora maggiore, per i ciborg derivanti dall'ibridazione di esseri umani con manufatti artificiali (si pensi al poliziotto ciborganico del film *Robocop*, cui non si possono non attribuire ricordi, sentimenti e strazi affatto umani). Il ciborg merita affetto e compassione oppure è uscito definitivamente dal consorzio umano per entrare in una sfera vaga e indefinibile e diventare preda di cacciatori senza scrupoli? I replicanti di *Blade Runner*, splendori di androidi e andreidi di dubbio statuto, debbono proprio essere eliminati? Insomma: chi decide che cosa significa essere umano e averne la dignità? Forse bisognerà presto riscrivere una "Carta dei diritti" da estendere a esseri la cui definizione sfugge per il momento ad ogni tentativo classificatorio.

Si rifletta anche che lo struggente desiderio che i robot o gli androidi o i ciborg manifestano di diventare del tutto umani sulla base di un consapevole "senso di inferiorità", desiderio che diviene ossessivo in Pinocchio e addirittura grottesco nel film di Spielberg, (AI): *Intelligenza Artificiale*, è frutto al solito di una nostra proiezione. Che motivo avrebbero creature tanto diverse da noi (e forse tanto migliori di noi) per voler diventare proprio come noi, se non quello di compiacere i loro vanitosi creatori? Ancora una volta i desideri dei genitori vengono proiettati sui figli con conseguenze forse disastrose.

A questo proposito, come ho accennato, alcuni ritengono che un giorno si potranno costruire robot più buoni degli esseri umani in virtù di un processo evolutivo che, innescato da noi, procederebbe poi in modo svincolato dai nostri condizionamenti. In fondo se noi siamo, in molte circostanze, aggressivi e malvagi ciò è dovuto al valore di sopravvivenza che queste caratteristiche hanno avuto nel corso dell'evoluzione. Ma i robot si evolveranno in un ambiente molto diverso dal nostro: l'ambiente dei robot, in gran parte, *siamo noi*. Ecco perché, si pensi al caso dei robot soldato, se vogliamo che questa nuova stirpe sia migliore di noi e magari ci aiuti a migliorare noi stessi (perché l'ambiente dell'uomo potrebbero un

### Robot e ciborg

La costruzione dell'uomo artificiale può seguire due strade, quella che porta al robot e quella che porta alle creature ciborganiche. In altri termini: o imparare dalla natura e imitarla (robot), oppure interferire con la natura e modificarla (ciborg). Nei robot confluiscono, si fondono e si unificano tre categorie di protesi:

- le protesi motorie e attive: le macchine semplici, le pinze, le automobili, i razzi ecc.;
- le protesi percettive: gli occhiali, gli sfigmomanometri, i microscopi, i nasi artificiali, le telecamere ecc.;
- le protesi cognitive: la scrittura, la matematica, le biblioteche, il computer, l'intelligenza artificiale ecc..

Il robot inoltre è caratterizzato da un certo grado di autonomia e da una certa capacità di apprendimento, che lo rendono un candidato plausibile a un'evoluzione corpo-mentale di tipo sia umanoide sia alternativo all'umano. L'evoluzione imitativa dell'umano potrebbe portare a macchine indistinguibili dall'uomo per le funzioni (intellettuali, attive, percettive, emotive) anche se distinguibili per i materiali e in parte per la struttura. Si tratta comunque di precisare i meccanismi dell'evoluzione, che appare eterodiretta e fortemente finalizzata, a differenza di quella biologica e, anche di quella culturale, che sono intrise di aleatorietà e contingenza.

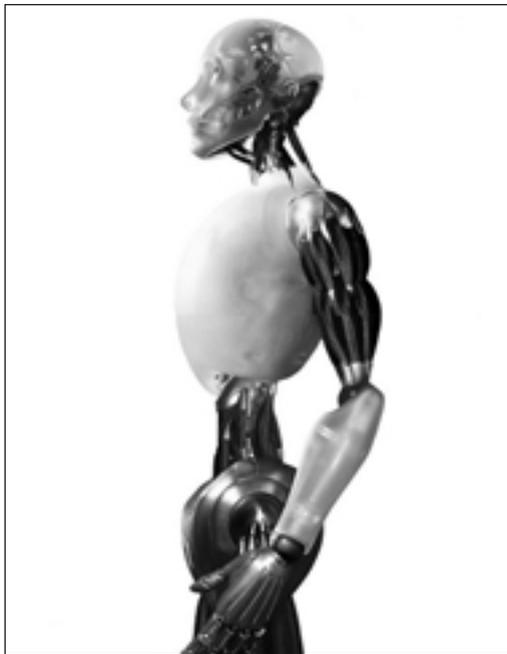
La convergenza di funzioni e strutture robotiche verso quelle umane prelude a una confusione tra naturale e artificiale. Ma più che nel robot questa confusione è evidente nei cyborg o ciborg, cioè nelle creature cibernetico-organiche derivanti da un'ibridazione spinta che, partendo dall'uomo, mira a sostituire parti sempre più ampie e complesse del corpo umano (o animale) con componenti di sintesi: braccia, mani, occhi, cervello ecc.. Ad un estremo di questo processo vi sono i trapianti d'organo, in cui l'ibridazione si mantiene sul piano organico-organico, all'altro estremo si colloca l'uomo artificiale, in cui non vi sono più residui organici e la sostituzione è completa. La spinta verso questa sostituzione progressiva deriva, almeno in parte, dalla consapevolezza che il corpo e le sue parti sono deteriorabili e quindi destinate a soccombere e a far soccombere il complesso di cui fanno parte. Il robot, per converso, parte da una base tutta artificiale e mira all'imitazione della funzione. Ma il punto d'arrivo appare lo stesso: l'uomo artificiale.

La confusione tra naturale e artificiale potrebbe prima o poi portare alla confusione tra umano e non umano e aprirebbe lo spinoso problema della definizione di persona: quali sono i "requisiti minimi" che un ente deve possedere per essere dichiarato persona e quindi avere la dignità corrispondente? Esiste un grado di imitazione funzionale o di sostituzione protetica al quale è lecito, o inevitabile, parlare di umanità, e quindi di dignità, dell'artefatto? Un'altra domanda che scaturisce da queste considerazioni: gli artefatti imitativi potrebbero indurre cambiamenti nella nostra concezione del corpo e della natura umana (così come l'intelligenza artificiale ha modificato la nostra concezione di intelligenza)? La costruzione dell'uomo artificiale potrebbe elevare gli artefatti a livello dell'uomo, oppure abbassare gli umani a livello delle macchine.

giorno essere *loro*) dovremmo stare molto attenti all'"indole artificiale" che imprimiamo in queste creature, pur nei limiti delle derive imprevedibili dovute alla loro autonomia. In questa prospettiva, instillare nei robot il desiderio di uguagliarci potrebbe segnare un regresso o almeno un ostacolo alla loro evoluzione etica verso la bontà (Figura 10). (Queste rapide considerazioni potrebbero e forse dovrebbero ampliarsi e dar luogo a una discussione approfondita sul "principio di precauzione" nell'ambito della roboetica).

### 10. LE EMOZIONI ARTIFICIALI

Le emozioni sono per noi umani un tratto costitutivo fondamentale, inseparabile dalle altre nostre caratteristiche. Le emozioni sono strettamente intrecciate alla razionalità computante, ma anche alle funzioni fisiologiche, alla memoria, all'esperienza, sono profondamente innestate nel corpo, inteso sia come insieme di organi sia come depositario della nostra identità e della nostra storia. Le emozioni sono tanto pervasive che ogni nostro atto si colora di esse e ogni nostra relazione con noi stessi e con l'"altro" ne è condizionata. Ma co-



**FIGURA 10**

Un robot del film  
*I, robot* di Alex  
Proyas (2004)

sa succede quando l'"altro" è inanimato, quando cioè non possiede emozioni da scambiare con le nostre? In questo caso facciamo tutto noi: investiamo l'oggetto di un'intensa proiezione affettiva e giungiamo al punto di attribuirgli proprietà che non possiede. Dietro

lo schermo di un computer immaginiamo una intelligenza (quasi) umana, dietro la condotta e gli atteggiamenti di un robot immaginiamo sentimenti, giudizio e consapevolezza.

Un esempio di questa proiezione-attribuzione affettiva è offerto dal robot cane Aibo, di cui la Sony ha di recente interrotto la produzione dopo averne costruito, dal 1999 al 2006, oltre 150.000 esemplari. Nel sito a lui dedicato, si legge che Aibo è un compagno gradevole e un intrattenitore nato, possiede l'istinto di girellare, cerca i suoi giocattoli e comunica col padrone, di cui riconosce la voce e il volto. Gli piace la musica e fa commenti sulle proprie sensazioni. Come per tanti robot, la personalità di Aibo si sviluppa tramite l'interazione con le persone e in base all'esperienza. Insomma un compagno affettuoso e discreto, che non ha bisogno di cibo, non sporca, non chiede di fare la passeggiatina e che si può disattivare quando non "serve": quanti vantaggi rispetto a un esigente e rumoroso cucciolo biologico!

Da tempo ormai alla compagnia di un animale

domestico si riconosce un notevole potere antidepressivo e ansiolitico, ma uno studio della Purdue University (Indiana, Stati Uniti) ha confermato che anche i robot zoomorfi possiedono queste doti. Su 72 bambini tra i sette e i quindici anni intervistati nell'indagine (tutti possessori di un Aibo) 50 hanno dichiarato che i robot sono buoni compagni. L'interazione con gli animali migliora il benessere psicologico dei bambini e la loro capacità di socializzare e di apprendere, ma ora il termine "animali" deve essere forse esteso a comprendere anche Aibo e i suoi colleghi, come Paro, un cucciolo robotico di foca, il celebre pulcino Tamagochi e altri ancora. I ricercatori sostengono che lo studio dei rapporti tra i bambini e gli zoorobottini mira a comprendere meglio lo sviluppo infantile e che nessuno ritiene che i robot sostituiranno mai gli animali; eppure in una società dove i rapporti umani sono sempre più rari e frettolosi la prospettiva di delegare alle macchine parte della nostra responsabilità comunicativa e affettiva non è poi tanto remota. Con quali conseguenze? È un tema da affrontare.

Un esempio reale di proiezione affettiva è fornito dal film *Grizzly Man* (Figura 11), di Werner Herzog, che narra la storia (vera) di Timothy Treadwell, un quarantenne disadattato che vive in Alasca a contatto con i temibili grizzly, trasgredendo, in uno slancio di empatia, il confine tra il sé e l'altro (*riquadro*). Proiettando sugli orsi il proprio amore ai limiti del morboso, addirittura illudendosi di identificarsi con loro, Timothy si illude di esserne ricambiato con lo stesso calore. Ma uno dei grizzly, che non gradisce questo travalicamento di confine e que-

#### FIGURA 11

Un'immagine tratta dal film *Grizzly Man*, di Werner Herzog (2005)



#### Il robot amoroso

Un tema molto particolare, affrontato ma non ben risolto nel film (AI): *Intelligenza Artificiale*, riguarda la costruzione di un robot che ci ami: è in un certo senso la situazione inversa rispetto a *Grizzly Man*, dove è l'uomo che ama l'orso, cioè l'alieno. Creando un robot che lo ami, il costruttore raggiunge un vertice di egocentrismo: infatti, non è previsto (o necessario) che l'uomo ricambi l'amore della creatura. È facile e insieme rischioso tracciare un parallelo con il rapporto tra Dio e uomo. Come dice il catechismo, Dio ci ha creati per conoscerlo e amarlo. Se poi Dio ci ami è un problema molto più complicato, spesso risolto in modo sbrigativo affermando che l'amore di Dio è testimoniato dal fatto che ci ha messi al mondo, come se vivere fosse in sé un bene, cosa su cui non tutti concordano. Certo, a livello umano, l'assenza di simmetria nel rapporto d'amore può condurre a situazioni molto dolorose, che molti uomini e donne conoscono per esperienza. Il robot amoroso ci pone di fronte al conflitto tra la consapevolezza di aver di fronte una macchina (non degna d'amore) e l'inequivocabile comportamento affettuoso della macchina, che unito al suo aspetto antropomorfo spinge alla proiezione emotiva e al coinvolgimento. Del resto anche nel rapporto amoroso tra umani la proiezione svolge un ruolo fondamentale: non ci s'innamora mai di una persona, ma dell'immagine che si ha (e che si costruisce) di quella persona. Comunque sia, se s'instaura un rapporto amoroso bilaterale, ne deriva per l'umano un'assunzione di responsabilità nei confronti dell'essere amato, anche se è una macchina. Come afferma Antoine de Saint-Exupéry nel *Piccolo Principe*, non c'è amore senza assunzione di responsabilità. Allora, come si esprimerebbe questa responsabilità nei confronti del robot amoroso? E che forme potrebbe rivestire l'amore per un robot, al di là del semplice rifiuto di considerarlo "solo" una macchina? Forse possiamo trarre qualche indicazione dal caso, meno perturbante e più realistico, dell'affetto-amore per gli animali domestici (si pensi anche a certe forme di feticismo).



sta promiscuità eccessiva, lo uccide e lo divora. Non voglio insinuare in alcun modo l'idea che il robot possa comportarsi in questo modo, ma non posso neppure escludere che dai robot attuali possano discendere, per evoluzione, creature aliene, così diverse da noi da non riconoscerci più ne come loro creatori e padroni e neppure come compagni da rispettare. Tornando all'attualità dei rapporti uomo-robot, il problema non riguarda solo i bambini: si pensi al numero crescente di anziani le cui famiglie non vogliono o non possono dedicare loro tempo e attenzione e che vengono accuditi da robot badanti. La possibilità di sostituire, almeno in parte, i rapporti umani con i rapporti robotici conferma la grande capacità di proiezione affettiva degli uomini, i quali tendono a interpretare azioni e reazioni puramente meccaniche (ma sono proprio tali? In altre parole: che cosa vuol dire "meccanico"?) come comportamenti intelligenti e coloriti di sentimenti: in fondo viviamo di apparenze. La cosa è preoccupante, poiché dimostra la capacità della tecnica di insinuarsi subdolamente in noi per strade insospettabili, creando forme di dipendenza e vere e proprie "zone di anestesia" nella nostra diffidenza e nel nostro distacco verso gli artefatti (Figura 12). Alcuni vedono in questa invasione progressiva una minaccia, tanto che in Giappone, Paese all'avanguardia nella robotica, si medita di non dotare i robot badanti di sembianze troppo umane, per evitare attaccamenti morbosi. Ma quando si parla di emozioni artificiali, si intende anche qualcosa che vada oltre la nostra proiezione: si stanno progettando agenti ca-

paci di *manifestare* emozioni (con l'espressione, con l'atteggiamento e così via) e, un domani, si vorrebbero costruire agenti capaci addirittura di *provare* (oltre che manifestare) emozioni. È un problema strettamente legato a quello della coscienza e porta a considerazioni dello stesso tipo. Si può dire che un agente artificiale manifesta emozioni quando si comporta in modi che, negli umani, presuppongono emozioni. Che poi si tratti di emozioni simulate, anche se riconoscibili per via comportamentistica (come nel criterio di Turing per l'intelligenza delle macchine), oppure di emozioni vere, di tipo psicologico e riflesse nella coscienza, resta un problema aperto e molto arduo (*riquadro*).



**FIGURA 12**

Il cosiddetto "Turco", un automa scacchista costruito nel 1770 dal barone ungherese Wolfgang von Kempelen. In realtà, a quanto pare, la base dell'automata celava un nano di grande abilità nel gioco

### La coscienza artificiale

A proposito degli artefatti più avanzati, come i robot, si è cominciato a parlare di *coscienza artificiale*. L'intelligenza artificiale riguarda attività che se compiute da un umano richiederebbero intelligenza, analogamente si può parlare di coscienza artificiale con riferimento ad attività che, se compiute da un umano, richiederebbero coscienza. È chiaro che qui "coscienza" significa consapevolezza e non coscienza morale (come nelle locuzioni: mi rimorde la coscienza, si metta una mano sulla coscienza e così via). Il problema centrale nel dibattito che si è avviato è se un robot possa, in linea di principio, manifestare una vera coscienza, nel senso psicologico, cioè una coscienza "in senso forte", oppure una semplice coscienza funzionale, o simulata, una coscienza "in senso debole". Il problema ha una grande rilevanza etica, poiché tutti i nostri comportamenti significativi sotto il profilo etico presuppongono la coscienza. È ormai evidente che esistono agenti dotati di capacità cognitive che non posseggono affatto coscienza (per esempio i programmi che giocano a scacchi), ma certe attività cognitive (umane) sembrano richiedere la coscienza.

La costruzione di enti dotati di coscienza "in senso forte" aprirebbe una serie di problemi etici: a tali enti dovrebbe essere riconosciuta una dignità analoga alla nostra ed essi avrebbero nei confronti nostri e di altri agenti quella responsabilità che nasce dalla consapevolezza dei propri atti. La coscienza potrebbe indurre in questi enti una certa capacità di soffrire, e a noi imporrebbe nei loro confronti un comportamento etico, che escluderebbe lo schiavismo e i maltrattamenti.

Alcuni ricercatori ritengono possibile la costruzione di agenti con una coscienza in senso forte, altri sono scettici, altri ancora addirittura contrari a questa prospettiva. Comunque sia, almeno in linea di principio il problema della coscienza artificiale si intreccia con molti dei temi trattati, in particolare con i concetti etici che concernono le Leggi di Asimov: il problema del bene e del male, del danno e dell'autodifesa e così via.



**FIGURA 13**  
 Il robot umanoide Asimo (*Advanced Step in Innovation Mobility*) della Honda si presenta a una signorina

## 11. IL SENSO

Se alcuni temono le proiezioni, gli equivoci emotivi e le confusioni di ruolo tra umani e robot, altri invece propendono per una visione in cui la tecnica contribuisce a una crescente apertura dell'uomo grazie al dialogo con l'alterità che si realizza mediante una connessione sempre più estesa tanto sul piano spirituale, cognitivo ed emotivo quanto su quello concreto. In questa prospettiva, esaltata dalla tecnologia, l'uomo coinvolge nella sua attività conversativa e dialogica tutta la realtà materiale, naturale e artificiale, e ogni oggetto contribuisce, attraverso l'uomo, a una progressiva crescita di significato, o meglio di "senso". A questo proposito, le azioni macchiniche, per quanto raffinate, ci appaiono comunque, almeno per il momento, prive di "senso", o meglio hanno senso per noi ma non per i robot. Il senso delle *nostre* azioni non sta nelle azioni, ma le precede, sta nel contesto e nella storia, negli affetti, nella gioia, nella speranza, nel dolore, nell'anticipazione. Si pensi al vasto territorio del simbolico, all'attività artistica, all'anelito verso lo spirituale e il trascendente. Le azioni delle macchine, per lo-

ro, non hanno senso perché la loro storia e il loro contesto siamo noi. È come se le macchine recitassero una poesia in una lingua a loro sconosciuta, ma che noi comprendiamo benissimo (Figura 13). Sono sempre gli uomini che interpretano ciò che le macchine fanno. Almeno per ora.

Ma evidentemente la proiezione emotiva sul robot umanoide non ci basta: come ho accennato, i ricercatori tentano di iniettare le emozioni nel robot stesso, per farne un vero e proprio interlocutore affettivo. Anzi, vanno ancora più in là: cercano di dotare queste macchine di una *coscienza*.

Su questa strada di umanizzazione profonda i problemi sono molti: in primo luogo non sappiamo che cosa sia la coscienza e non sappiamo come funzioni. Inoltre nell'uomo, emozioni, coscienza, razionalità, corporeità e quant'altro sono talmente intrecciate da rendere poco plausibile il procedimento seguito per dotarne i robot, che è di tipo additivo: a una base cognitiva di IA (*Intelligenza Artificiale*) si aggiunge un corpo (percezione artificiale ed esecuzione di funzioni), poi a questo complesso si aggiungono (come?) emozioni artificiali e poi, in cima a tutto, si deposita una coscienza artificiale.

Qui il termine "artificiale" indica la derivazione da processi diversi da quelli biologico-evolutivi e qualifica in modo essenziale i sostantivi ai quali si applica. Consideriamo l'IA, il cui scopo primo, benché non sempre dichiarato, è quello di replicare l'intelligenza umana: ebbene, i risultati sono caratterizzati molto più dall'aggettivo "artificiale" che dal sostantivo "intelligenza". L'IA è sì interessante, ma è radicalmente diversa dalla *nostra* intelligenza, e sarebbe opportuno adottare una terminologia altrettanto diversa. A scanso di equivoci e derive metaforiche fuorvianti converrebbe evitare termini molto impegnati come intelligenza, emozioni, coscienza.

## 12. IL DEMIURGO ALLO SPECCHIO

*Helena: perché li fabbricate, allora?*

*Busman: ahahah! Questa è bella! Perché si fabbricano i Robot!*

*Fabry: per il lavoro, signorina. Un Robot sostituisce due operai e mezzo. La macchina*

*umana, signorina, era molto imperfetta. Un giorno occorreva eliminarla definitivamente.*

Karel Čapek, *R.U.R.*

Un paio di osservazioni conclusive. La marcia sempre più rapida di una tecnologia raffinata e suggestiva come la robotica non può non avere effetti profondi sull'immagine che abbiamo di noi stessi e sul nostro stesso essere umani: specchiandoci in quello straniante *alter ego* che sta diventando il robot, quale immagine ci ritorna? L'impresa della robotica, cioè la costruzione di un vero e proprio *uomo artificiale*, potrebbe darci, per analogia o per contrasto, indicazioni utili su di noi, così come ha fatto l'IA. In questa prospettiva di rispecchiamento il robot potrebbe essere un laboratorio di etica (artificiale)?

Infine si pone la questione del perché: *perché costruiamo i robot?* In certi casi particolari la risposta è ovvia: per eseguire compiti pesanti o pericolosi o ripetitivi, oppure per sostituire la manodopera umana con vantaggio economico o di rendimento. Ma tutto ciò non risponde alla questione di fondo: perché costruire macchine così simili a noi? Qualche risposta possibile: l'umanità sta facendo di tutto per entrare nel novero delle specie estinte e, sentendo prossima la fine dell'avventura, vuole lasciare un segno della propria grandezza, perciò costruisce macchine che possano sopravvivere e che ricordino a chi verrà (chi? Le macchine stesse?) un passato di gloria. C'è anche, come si è detto, l'orgoglio tutto umano di forzare e imitare i segreti della natura. Da ultimo c'è lo scopo comune a tutte le forme d'arte e di tecnica: stupire. "E' del poeta il fin la meraviglia", cantava Marino, e Leonardo annotò: "farò una finzione che significherà cose grandi".

Qualunque risposta diamo alla domanda di fondo, "perché?", è indubbio che da essa scaturiscono subito altre questioni che ne mettono in luce la natura socioculturale ed etica: quale società vogliamo costruire progettando i robot? Quali valori cerchiamo di rafforzare o di indebolire? Molti ricercatori non dimostrano alcun interesse per questi problemi e procedono tranquilli o entusiasti sulla strada dell'innovazione tecnica. Altri si



**FIGURA 14**

*Nel 2003 il robot umanoide Kiro, della Sony, ha diretto un concerto di musica classica a Tokyo*

pongono in una prospettiva di breve respiro, conformandosi a codici simili alle leggi di Asimov. Altri ancora, una minoranza, si pongono nella prospettiva di medio e lungo termine e cercano di immaginare gli sbocchi possibili di quella che ormai è una vera e propria invasione dei robot. Qui le implicazioni della robotica e della roboetica si confondono con gli scenari elaborati in quell'attrezzatissima palestra di ipotesi sul futuro che è la fantascienza.

Il 13 marzo 2004, davanti a un folto pubblico di giovanissimi, l'orchestra filarmonica di Tokyo ha eseguito la *Quinta* di Beethoven sotto la direzione di KIRO, un robot umanoide della Sony, che, dopo qualche incertezza, ha fatto una discreta figura, aggiungendo un altro tassello al vasto mosaico delle attività umane eseguite (o imitate) dalle macchine (Figura 14). Ora, tanto per fare un esercizio di fantasiologia, mi immagino un nipotino di Kiro che dirige un'orchestra di robot davanti a un pubblico di robot: se venissero a mancare gli umani chi si porrebbe le questioni di cui stiamo parlando? Dove andrebbe a finire il problema del senso? Chi si chiederebbe che cosa? E infine: dove andrebbe a finire la follia degli uomini? Che fine farebbero l'arte, l'umorismo, la trasgressione, la creatività, il gioco, il nonsenso? Chi potrebbe avvertire la differenza tra una lacrima e una goccia di pioggia? Forse, come ho detto sopra, per perpetuare la follia creativa dell'uomo, ci sarebbe bisogno di una macchina schizofrenica. Ma chi saprebbe costruirla, e chi, sapendola costruire, se ne assumerebbe la responsabilità?

## Bibliografia

- [1] Bateson G.: *Verso un'ecologia della mente*. Adelphi, Milano, 2 edizione, 2000.
- [2] Buttazzo G.: Coscienza artificiale: missione impossibile?. *Mondo Digitale*, n. 1, marzo 2002.
- [3] Carlucci Aiello L., Dapor, M.: Intelligenza Artificiale: i primi 50 anni. *Mondo Digitale*, n. 10, giugno 2004.
- [4] Fukuyama F.: *L'uomo oltre l'uomo*. Mondadori, Milano, 2002.
- [5] Galván J.M.: *La robotica come speranza: la technoetica*. In: *La sfida del post-umano*, a cura di Sanna I., Ed. Studium, Roma, 2005.
- [6] Longo G.O.: *Il simbiote: prove di umanità futura*. Meltemi, Roma, 2003.
- [7] Longo G.O.: Uomo e tecnologia: una simbiosi problematica. *Mondo Digitale*, n. 14, giugno 2005.
- [8] Losano M.G.: *Storie di automi*. Einaudi, Torino, 1990.
- [9] Monopoli A.: *Roboetica*. (<http://www.roboetica.it/page2.html>).
- [10] Saint-Exupéry A.: *Il Piccolo Principe*. Bompiani, Milano, 2000.
- [11] Shelley M.: *Frankenstein, ovvero il moderno Prometeo*. Mondadori, Milano, 1983.
- [12] Veruggio G.: *La roboetica e le sfide della rivoluzione robotica*. In: *La sfida del post-umano*, a cura di Sanna I., Ed. Studium, Roma, 2005.
- [13] Veruggio G.: Il cammino della roboetica. *Le Scienze*, n. 461, gennaio 2007.
- [14] Wiener N.: *The human Use of Human Beings*. Cybernetics and Society, Avon Books, New York, 1967.
- [15] Zaccaria R.: Aspettando robot. *Mondo Digitale*, n. 7, settembre 2003.

## Filmografia

- 2001: *Odissea nello spazio*, regia di Stanley Kubrick, Stati Uniti-Gran Bretagna, 1968.
- AI: Intelligenza Artificiale*, regia di Steven Spielberg, Stati Uniti, 2001.
- Blade Runner*, regia di Ridley Scott, Stati Uniti, 1982.
- Grizzly Man*, regia di Werner Herzog, Stati Uniti, 2005.

GIUSEPPE O. LONGO è ordinario di Teoria dell'informazione nella Facoltà d'Ingegneria dell'Università di Trieste. Si occupa di codifica di sorgente e di codici algebrici. Ha diretto il settore "Linguaggi" del Laboratorio della "International School for Advanced Studies" (Sissa) di Trieste e il Dipartimento di Informazione del "Centre Internationale des Sciences Mécaniques" (Cism) di Udine. Socio di vari Istituti e Accademie, s'interessa di epistemologia, di intelligenza artificiale e del rapporto uomo-tecnologia. È traduttore, collabora con il Corriere della Sera, con Avvenire e con numerose riviste. È autore di romanzi, racconti e opere teatrali tradotti in molte lingue.

E-mail: [longo@univ.trieste.it](mailto:longo@univ.trieste.it)