

COSCIENZA E INTELLIGENZA ARTIFICIALE: LA COSCIENZA NEL DIALOGO TRA LE INTELLIGENZE

CONSCIOUSNESS AND ARTIFICIAL INTELLIGENCE:
CONSCIOUSNESS IN THE DIALOGUE BETWEEN INTELLIGENCES

GIORGIO BONACCORSO ¹

La coscienza è l'evidenza meno evidente: evidente perché tutti, in un qualche modo, ne facciamo esperienza; meno evidente perché si fa ancora tanta fatica a esaminarla col rigore scientifico con cui si affrontano tanti altri fenomeni dell'universo, della biosfera e dell'uomo. Come intendere o spiegare il fenomeno tanto impalpabile e contemporaneamente ineludibile della coscienza? Si può ben dire che la coscienza è stata contesa tra la filosofia, la scienza, l'arte, la religione, con risultati che non soddisfano mai tutti. Attualmente sono emersi due grandi contendenti, ossia le scienze che si occupano dell'intelligenza naturale e le scienze che si occupano dell'intelligenza artificiale. I punti di contratto tra neuroscienze e informatica sono molteplici anche se rimangono aperti ancora alcuni varchi difficilmente valicabili. In ogni caso, almeno per quanto riguarda la coscienza, non ci si può occupare dell'intelligenza artificiale prescindendo dalle neuroscienze, ossia dagli studi sull'intelligenza naturale, biologica.

1. L'intelligenza naturale: il pre-umano verso l'umano

La scienza, come la fisica e la biologia, studia le origini dell'universo e della vita e quindi si occupa di quel pre-umano senza il quale non potrebbe esistere l'umano. Sotto questo profilo oggi stanno emergendo sempre più le neuroscienze, ossia l'insieme delle scienze che si occupano del cervello e delle sue capacità mentali nel contesto della vita sulla Terra. Non si può comprendere l'intelligenza naturale, senza prendere le mosse dal lungo percorso evolutivo che ha portato all'*homo sapiens*, e che affonda le sue radici non solo nelle più antiche forme di sistemi neurologici ma ancora prima, ossia nel modo stesso di comportarsi di tutti gli organismi. Dal punto di vista biologico, la nostra specie è l'ultima erede di una lunga storia della vita sulla Terra: "ultima" almeno nel senso che più di ogni altra ha trasformato l'ecosistema del nostro pianeta. In questo lungo percorso si

RIASSUNTO

La coscienza umana si trova al crocevia tra biologia e tecnologia. Per un verso la coscienza si è andata formando nel lungo processo evolutivo che ha preceduto e realizzato l'emergere dell'*homo sapiens* e delle sue capacità emotive e cognitive (dal pre-umano all'umano). Per un altro verso, nei tempi più recenti la coscienza è stata arricchita e viene sempre più arricchita da supporti tecnici che ne possono espandere le capacità ma che ne possono anche stabilire il superamento per l'emergere di nuove forme di intelligenza (dall'umano al post-umano). L'attenzione al paradigma della complessità, evitando facili e ingenui riduzionismi, consente un'integrazione più appropriata di questi due processi.

Parole chiave

Corpo, cervello, evoluzione, computer, informatica, complessità.

SUMMARY

Human consciousness is at the crossroads of biology and technology. On the one hand, consciousness has been forming in the long evolutionary process that preceded and brought about the emergence of *homo sapiens* and our emotional and cognitive capacities (from the pre-human to the human). On the other hand, in more recent times consciousness has been and is increasingly enriched by technical support systems that can expand these capacities, but which can also lay the groundwork for the emergence of new forms of intelligence (from the human to the post-human). Focusing attention on the paradigm of complexity, while avoiding facile and naïve reductionism, allows a more appropriate integration of these two processes.

Keywords

Body, brain, evolution, computer, computer science, complexity.

inserisce quell'esperienza che riassumiamo col termine "coscienza".

1.1. La cognizione come fenomeno biologico

Lo studio degli organismi viventi porta a considerare il fenomeno dell'intelligenza che nel suo senso più originario riguarda la capacità di adattamento all'ambiente. Sotto questo profilo, l'intelligenza non aspetta la nascita dei neuroni per

manifestarsi, ma quando compare il sistema neurale, emergono livelli straordinari dell'intelligenza. Si tratta allora di comprendere come si possa misurare o comunque riconoscere un comportamento intelligente. Per molto tempo ci si è mossi tra chi ha insistito soprattutto sulle strutture biologiche che determinano i comportamenti degli animali, e chi ha insistito piuttosto sull'importanza dell'apprendimento. In tempi più re-

RESUMEN

La conciencia humana se encuentra en el cruce entre la biología y la tecnología. Por un lado, la conciencia se ha ido formando en el largo proceso evolutivo que precedió y provocó el surgimiento del homo sapiens y de sus capacidades emocionales y cognitivas (del prehumano al humano). Por otro lado, en tiempos más recientes la conciencia se ha enriquecido y se enriquece cada vez más con soportes técnicos que pueden ampliar sus capacidades pero que también pueden establecer su superación para el surgimiento de nuevas formas de inteligencia (de humana a poshumana). La atención al paradigma de la complejidad, evitando reduccionismos fáciles e ingenuos, permite una integración más adecuada de estos dos procesos.

Palabras clave

Cuerpo, cerebro, evolución, ordenador, informática, complejidad.

centi è emersa una spiegazione di tipo cognitivo. L'apprendimento è indubbiamente molto rilevante, ma consiste solo nell'accumulare un numero così enorme di informazioni che finirebbe per disorientare qualsiasi animale. La cognizione riguarda la capacità di selezionare queste informazioni in modo da trarne un vantaggio per l'adattamento.

Le capacità cognitive svolgono il loro prezioso compito relativamente alle

specificità dell'organismo che vi fa ricorso, con la conseguenza che sarebbe piuttosto ingenuo, misurare l'intelligenza degli animali sulla base della nostra intelligenza. Si dovrebbe, piuttosto, spostare l'attenzione e chiedersi, come appare già nel titolo di un libro, se noi uomini «siamo così intelligenti da capire l'intelligenza degli animali».² Indubbiamente, tante specie hanno scopi simili ai nostri, come procurarsi il cibo e avere la prole, ma le modalità con cui raggiungono questi scopi possono essere diversi dai nostri per il modo diverso di percepire le cose. Quando si vuole indagare l'intelligenza di un animale occorre ricorrere a test adatti alle caratteristiche dell'animale in questione. È piuttosto affrettato, per esempio, affermare che a differenza delle scimmie, gli elefanti non siano in grado di usare utensili. Secondo alcuni esperimenti, un elefante non sarebbe in grado di fare ciò che è normale per una scimmia, ossia ricorrere a un bastone per raccogliere una banana. Bisogna tener presente, però, che mentre la scimmia ricorre alle mani, l'elefante dispone solo della proboscide che gli serve tanto per fiutare il cibo quanto per afferrarlo. E così se afferra un bastone questo compromette la sua capacità di fiutare il cibo: per questo motivo non lo utilizza per prendere un cibo qualora questo fosse troppo lontano per afferrarlo direttamente con la proboscide. Se, però, il cibo è molto in alto e si mette a sua disposizione una cassa di legno abbastanza alta e ro-

busta, l'elefante sposta la cassa sotto il cibo, vi sale con le zampe anteriori e afferra il cibo con la proboscide.³ L'intelligenza si misura correttamente solo a condizione di considerare le caratteristiche specifiche, e anzitutto anatomiche, degli animali,⁴ con la conseguenza che non ci si può limitare al cervello ma si devono tener presenti anche altri fattori e in primo luogo il corpo in cui è sedimentato il cervello. Lo studio dei primati come pure di tanti mammiferi, mette in evidenza, per esempio, che la grandezza, il peso e il numero di neuroni di un cervello non costituiscono l'unico criterio per valutare l'intelligenza di un animale, né sono sufficienti a rendere ragione delle capacità specifiche dell'uomo⁵. Una delle acquisizioni più importanti delle neuroscienze è che il cervello non è una fonte autonoma delle funzioni mentali, ma è ciò di cui si serve il corpo per elaborare le funzioni mentali. Il fatto fondamentale è che il corpo costituisce quel contatto continuo con l'ambiente che consente al cervello di raccogliere le informazioni grazie alle quali il sistema neurale si modifica ed evolve. Tutto porta verso il radicamento corporeo, e non solo cerebrale, delle capacità cognitive, ossia a una sorte di «cognizione incarnata».⁶

1.2. La coscienza come esperienza somatica

Nel quadro appena descritto si inserisce anche la coscienza, ossia quell'esperienza piuttosto difficile da definire che comunque, in un modo o

nell'altro, consente a un individuo di riconoscersi. Un esperimento classico per verificare l'esistenza di questa esperienza in un determinato individuo è quello dello specchio, come quando un elefante su cui è stato disegnata una croce sopra un occhio specchiandosi tenta di cancellarla o quando uno scimpanzé si specchia di spalle per vedere la sua parte posteriore. Questi e altri esperimenti mostrano che molti animali hanno un qualche grado di coscienza.⁷ L'aspetto più rilevante è che la coscienza è un fenomeno biologico che emerge prima dello sviluppo di una razionalità concettuale astratta come si può osservare in un uomo adulto: «Anche se noi non possiamo misurare direttamente la coscienza, altre specie presentano precisamente le capacità tradizionalmente considerate come suoi indicatori».⁸

Un indicatore fondamentale è costituito dai dati neurobiologici e comportamentali che riguardano la sfera emotiva. Tra le principali caratteristiche degli organismi viventi troviamo l'azione, ossia il movimento intenzionato. Ma come si forma l'intenzione? In termini molto sintetici si può affermare subito che la componente intenzionale dell'azione è costruita anzitutto dall'emozione:⁹ proprio in questa spinta emotiva dell'azione si elaborano le basi cognitive degli esseri viventi.¹⁰ Per comprenderlo fino in fondo occorre liberarsi da pregiudizi costruiti nei secoli e che penalizzano tutto ciò che attiene alla sfera emotiva. Anzitutto, le emo-

zioni sono riconoscibili soprattutto nelle espressioni facciali che, fin dagli studi di Ekman, risultano essere universali. Inoltre il ricorso alle espressioni facciali per esprimere e comunicare emozioni è rilevante in molti primati anche se con specificità non trascurabili. In molte specie, quindi, le emozioni sono reazioni che realizzano comunicazioni importanti con l'ambiente e con gli altri. Com'è stato affermato, in molti casi, le emozioni «sanno meglio di noi che cosa ci fa bene, anche se spesso non siamo pronti ad ascoltare». ¹¹ Va da sé, che quel «meglio di noi» si riferisce al noi razionale. Le emozioni «sono lo strumento con cui ci assicuriamo di fare quello che è meglio per noi. Inoltre soltanto il corpo può compiere le azioni richieste.

La mente di per sé è inutile: ha bisogno del corpo per interagire con il mondo. Le emozioni sono collocate all'intersezione di tre entità: mente, corpo e ambiente». ¹²

Si pensi alla «teoria della mente», con cui si intende la capacità di comprendere gli stati mentali degli altri. La stessa terminologia (teoria della mente) sembra poco adatta perché, com'è stato segnalato, troppo cerebrale e disincarnata. «Gran parte della nostra conoscenza di noi stessi viene dall'interno del nostro corpo, e gran parte di ciò che sappiamo degli altri proviene dalla lettura che facciamo del loro linguaggio del corpo». ¹³ Alla base non sta la lettura della mente ma la lettura del corpo, non la teoria della mente ma l'empatia. ¹⁴ «Proviamo

gli stessi sentimenti degli altri adottando a nostra volta le loro posture, i movimenti e le espressioni. L'empatia si trasmette da un corpo all'altro». ¹⁵

Il discorso sull'empatia ci introduce nel tema del rapporto intersoggettivo, che è quanto dire intercorporeo. Un aspetto decisivo per la coscienza. La dinamica tra azione ed emozione è decisivo ma pervenire alla coscienza occorre anzitutto che l'azione sia riconosciuta come opera di un Sé. Non sembra possibile, però, che ciò si verifichi senza l'altro, o meglio senza vedere l'azione dell'altro. Com'è stato osservato, lo stretto legame tra l'autoidentificazione e il riconoscimento delle proprie azioni non dipende solo dall'esecuzione di un movimento ma anche dal vederlo. ¹⁶ La coscienza affonda le sue radici nell'azione secondo il duplice versante del compiere l'azione e del veder compiere l'azione. ¹⁷

La dimensione intersoggettiva appare in tutta la sua rilevanza quando si tenta di capire le scelte volontarie e libere. L'intento di tanti studi è di localizzare la zona cerebrale che è alla base della libera volontà, chiedendo, per esempio, a un soggetto di alzare un dito quando viene toccato, e verificare così cosa accade nel sistema nervoso centrale. Dalle ricerche risulta l'aumento di attività della corteccia prefrontale. ¹⁸ Secondo alcuni, però, quando si fanno gli esperimenti accade qualcosa su cui è opportuno riflettere. Infatti, il soggetto sottoposto all'esperimento non può che «eseguire liberamente una risposta preselezio-

nata. Che razza di libertà è mai questa». ¹⁹ Anche quando, si chiede ai soggetti di alzare il dito quando sentono l'urgenza di farlo, «devono ridurre drasticamente la loro libera volontà. Hanno bisogno di dare istruzioni a se stessi comportandosi più o meno in questo modo: "Dovrò fare in modo che ci sia un intervallo ogni volta diverso, ma non troppo, tra il sollevamento di un dito e il successivo, in modo che lo sperimentatore non possa predire facilmente quando solleverò nuovamente il dito". I soggetti, in realtà, non stanno facendo libere scelte sulle loro azioni; stanno giocando una complessa partita con lo sperimentatore». ²⁰ Ma allora da dove proviene la libertà: dalla corteccia prefrontale o dall'intreccio con un altro, ossia con lo sperimentatore? Dove risiede l'io che sceglie in piena coscienza e libertà? Nella corteccia prefrontale? Il primato accordato alla corteccia prefrontale viene dal fatto che si pensa a un individuo isolato, o se si vuole a un cervello isolato, trascurando che tutto avviene entro l'intreccio di più individui e più cervelli. «Il cervello umano è sintonizzato in maniera assai fine per le interazioni con gli altri. Concetti come volontà, responsabilità e significato emergono da queste interazioni». ²¹ Interazioni che coinvolgono ovviamente tutto il corpo e in modo particolare le emozioni, per cui se si vuole individuare l'io non si può pensare a un omuncolo nascosto in qualche area cerebrale. «Anziché una singola area che fa

delle scelte, forse c'è una rete di aree che applicano dei vincoli per determinare la scelta finale. Tali vincoli provengono da più fonti: i nostri corpi (ci sono alcune azioni che è fisicamente impossibile eseguire); le nostre emozioni (ci sono azioni che potremmo rimpiangere). E soprattutto, ci sono vincoli che provengono dal mondo sociale», ²² dato che ci sono azioni che non si compiono di fronte a chiunque. Il primo vincolo, il corpo, è anche il luogo complesso grazie al quale il cervello esiste e opera: soprattutto è il luogo delle emozioni come reazioni interne a ciò che avviene fuori e il luogo dell'empatia come modalità di intrecciare quei rapporti intersoggettivi che costituiscono la condizione di possibilità della vita sociale, come il ricorso a una tecnica condivisa dal gruppo, l'elaborazione di comportamenti morali, le espressioni artistiche, le credenze religiose.

La coscienza rimane comunque un terreno ancora tutto da esplorare. La difficoltà più evidente è che si tratta della più intima esperienza soggettiva, e quindi un fenomeno che espone a sforzi enormi la scienza il cui intento primario è di giungere a conclusioni oggettive. Ciò che comunque emerge da tante ricerche e a cui si è fatto riferimento nelle pagine precedenti è l'intreccio tra molteplici fattori e in primo luogo dell'emozione e dell'intersoggettività. Diversi studiosi, però, insistono sulla dimensione computazionale. Alla base sta la convinzione che il comportamento più originario dell'universo sia l'informazione. In ef-

fetti l'energia e quello stato dell'energia che è la materia sono rette da un'intensissima rete di informazioni a partire dallo scambio di messaggi tra le particelle elementari, gli atomi e le molecole. La vita, così come può essere osservata sul nostro pianeta, è retta da particolari modalità di scambio di messaggi, che sono le informazioni interne ed esterne delle proteine e dei virus, delle cellule e delle molecole del DNA, dei neuroni e delle trasmissioni sinaptiche elettriche e chimiche, come pure tra gli organi di un'unità corporea come i rettili, i mammiferi, i primati, e tra i primati l'uomo. Anche le azioni, le emozioni, le capacità cognitive, gli scambi sociali e infine anche la coscienza, possono essere considerate dal punto di vista informatico e computazionale. Ed è questo punto di vista che svolge un ruolo decisivo nella ricerca tecnologica di macchine viventi e intelligenti, dato che è il punto di vista che, secondo tanti, costituisce la via privilegiata per un confronto tra l'intelligenza naturale e l'intelligenza artificiale. Nell'affrontare alcuni tentativi di elaborazione dell'intelligenza artificiale sarà quindi opportuno tornare agli studi neurobiologici che danno ampio spazio all'aspetto computazionale delle capacità cognitive dell'uomo e della coscienza.

2. L'intelligenza artificiale: l'umano verso il post-umano

L'affermazione secondo cui la nostra specie è l'ultima erede dell'evoluzione della vita sulla Terra, secondo alcuni

deve essere corretta, dato che potrebbe essere la penultima erede, a cui succederebbe l'intelligenza artificiale. Alla base sta la tecnica che, nelle sue diverse forme e livelli, tende a trasformare il mondo e lo stesso uomo. Riguardo all'uomo la tecnica tende sia a modificarne le componenti biologiche anche con protesi sempre più sofisticate, sia a realizzare supporti e programmi sempre più avanzati, ossia computer e robot in grado di svolgere molte delle attività umane e di superarne le potenzialità. Tutto ciò porterebbe a una sorta di passaggio dall'umano al post-umano.

2.1. La condizione di neuroplasticità

L'attuale epoca geologica, ossia l'olocene (iniziato poco più di 11.000 anni fa), ha visto un tale incremento dell'influsso dell'*homo sapiens* (comparso almeno 200.000 anni fa) sull'ambiente che si può arrivare a parlare di antropocene. L'estensione di quell'influsso è particolarmente evidente in questi ultimi decenni, nei quali le capacità tecniche dell'uomo si sono indirizzate in modo sempre più considerevole verso la costruzione di un'intelligenza programmata su basi non biologiche, o almeno diverse da quelle tipicamente biologiche dell'essere umano. Si può intendere l'epoca in cui viviamo come una «collaborazione sempre più stretta fra la nostra eredità biologica e un futuro che trascende la biologia».²³ Un'epoca che vede in gioco la stessa comprensione dell'essere umano. Se per secoli sono state

soprattutto le religioni e le filosofie a tentare di individuare la natura dell'uomo, in tempi più recenti sono state le scienze e in particolare discipline come l'antropologia culturale e la psicologia. Secondo alcuni, però, queste due scienze avrebbero dei limiti superati dalle scienze della mente che consentono di pervenire a una nuova sintesi.²⁴ Lo studio delle basi biologiche della mente oltre a far luce sulla natura umana, aprirebbe anche le porte ai tentativi di costruire macchine intelligenti.

La possibilità che la costruzione di macchine intelligenti influisca sull'essere umano e in particolare sulle sue capacità cognitive, dipende anzitutto dal fatto che il sistema neurale è molto elastico e pronto a modificarsi nel contatto con i fenomeni ambientali e interni. Gli studi di questi ultimi decenni hanno messo in evidenza questa plasticità del cervello, mostrandone la rilevanza per l'evoluzione del cervello stesso. Se l'epoca in cui viviamo è quella della società liquida, la mente da sempre è liquida nel senso che i sistemi neurali che ne costituiscono la base sono molto fluidi. Per questo si è fatto ricorso a termini come «plasticità neuronale» o «neuroplasticità», con i quali si intendono i mutamenti della configurazione neurale del cervello nel tempo e che riguardano «sia i cambiamenti *strutturali*, riguardanti l'insieme dei neuroni e le loro connessioni, che quelli *funzionali*, relativi al comportamento del singolo neurone [...]». Da un punto di vista cognitivo, conside-

rando le interazioni dell'uomo con gli stimoli esterni, la neuroplasticità si manifesta attraverso la capacità del cervello di riorganizzare la propria configurazione neuronale in maniera tale da adattarsi agli stimoli provenienti dall'ambiente e, di conseguenza, anche a quelli prodotti dalle macchine».²⁵

L'influsso delle macchine sulla mente umana non è legata solo alla plasticità del cervello ma questa costituisce un ambito sempre più decisivo nel rapporto tra l'intelligenza naturale e l'intelligenza artificiale. L'aspetto di non secondario interesse è che se per un verso la neuroplasticità mette l'intelligenza naturale nelle condizioni di subire l'influsso dell'intelligenza artificiale, per un altro verso ne definisce anche la differenza dell'intelligenza artificiale almeno fino a quando quest'ultima non avrà la stessa plasticità. L'ipotesi da verificare è se la plasticità del cervello umano dipenda dal fatto di non essere stato programmato da un altro cervello. Il tema centrale è costituito quindi dalla nozione di «programma» e da quanto margine alla plasticità lasci il software con cui si programma un computer.

2.2. Il programma informatico come base della cognizione naturale e artificiale

Si è detto che il cervello non è stato programmato da qualcuno come lo è la macchina, ma questo non significa che non si comporti secondo un programma o una serie di programmi. Ed è su questo punto che per molti si gioca il confronto tra intelligenza

naturale e intelligenza artificiale. Cos'è dunque un programma? Per rispondere occorre tener presente i presupposti a partire dei quali si tenta di affrontare il problema. Un presupposto è stato chiaramente esposto da M. Minsky, uno dei padri del lavoro sull'intelligenza artificiale. Per spiegare cosa sia l'intelligenza Minsky scrive: «Vogliamo spiegare l'intelligenza come una combinazione di cose più semplici». ²⁶ Sulla base di questa convinzione si può considerare la possibilità che a partire dalle cose semplici che costituiscono l'intelligenza biologica si possa costruire l'intelligenza artificiale, sia pure su differenti supporti fisici. Va notato subito, anche per quanto si dirà in seguito, che sebbene questo autore sottolinei l'enorme complessità dell'intelligenza, rifiuta l'idea che il tutto sia più della somma delle sue parti. ²⁷ La convinzione che *il tutto sia semplicemente la somma delle parti* comporta la possibilità di elaborare un programma in grado di organizzare un certo numero di componenti (le parti) e produrre un sistema complesso (il tutto). In questa prospettiva l'intelligenza artificiale è un programma mentale, un software, che può essere trasferito da un supporto a un altro, da un hardware a un altro, e anche, per esempio, da un supporto fisico di tipo biologico a un supporto fisico non biologico. La rilevanza di un programma è ineccepibile. Un computer, per esempio, è una combinazione binaria che opera sulla base delle quattro fondamentali funzioni logiche. Questa combinazione

è il programma di un sistema che rende possibile la cosiddetta intelligenza artificiale. Anche i sistemi esperti, ossia i sistemi capaci di imparare e di fornire le soluzioni migliori a un determinato problema, operano grazie a un programma. ²⁸

Il problema è di elaborare macchine con un'intelligenza paragonabile a quella biologica e soprattutto umana. ²⁹ Non si tratta di copiare il corpo, il cervello e la mente dell'uomo, così come per ottenere un aeroplano non bisogna costruire una macchina che copi fisicamente un uccello (con ali mobili e piume). ³⁰ Ciò non toglie che occorre avere delle macchine che, come l'essere umano, siano capaci di quella plasticità che rende possibile non solo la capacità di apprendere ma anche di modificarsi in ragione di quanto appreso. Potremmo dire che se in virtù della neuroplasticità la tecnologia può influire sul cervello, quella stessa tecnologia deve avere la medesima plasticità per raggiungere le capacità del cervello. Per questo si stanno estendendo gli studi per elaborare una intelligenza artificiale che sia biologicamente ispirata. ³¹

Il primo problema che emerge è costituito dal fatto che il cervello e quindi l'intelligenza naturale sono il risultato di un lungo percorso evolutivo. Se si può avere un'evoluzione artificiale, occorre tener presente le somiglianze ma anche le inevitabili differenze rispetto all'evoluzione naturale. La più evidente differenza è che l'evoluzione naturale non è programmata ma opera attraverso un

processo di adattamento con risultati non sempre prevedibili, mentre l'evoluzione artificiale si muove a partire da un programma col quale si tende a ottenere delle soluzioni a problemi predefiniti, con la conseguenza che la prima consenta una maggior creatività rispetto alla seconda.³² Se comunque si vuole procedere a costruire una macchina ispirata alla biologia, occorre tener presente che i pilastri dell'evoluzione biologica sono la popolazione, la diversità, l'ereditarietà, la selezione. È, per esempio, fondamentale un intreccio tra computer e robot per riprodurre la dinamica tra genotipo e fenotipo.

Una questione fondamentale è costituita dal fatto che la prima evidenza dell'evoluzione è il suo cammino verso la complessità. Ma che cos'è la complessità per una macchina? E si può progettare una macchina che sia in grado di evolvere verso sistemi sempre più complessi, così come avviene per un organismo vivente e il cervello in esso contenuto? Molto dipende dal rapporto, segnalato già con Minsky, tra le componenti (relativamente) semplici che sono alla base della vita e i sistemi complessi costituiti da quelle componenti semplici. Secondo Minsky e tanti altri, la vita ha raggiunto le capacità del cervello umano, perché è stata in grado di autorganizzarsi in modo sempre più complesso partendo da componenti più semplici. E così se a una macchina vengono forniti queste componenti essa potrà autorganizzarsi

ed evolvere fino a raggiungere, se non superare, le capacità mentali del cervello umano.³³ Alla base sta ovviamente il processo di adattamento ben presente nei «biological nervous systems» e che potrebbe essere riproducibile nell'«artificial neural networks».³⁴ L'ipotesi più probabile è che i due processi si intreccino e formino sistemi neurali ibridi.³⁵

Una intelligenza artificiale biologicamente ispirata deve affrontare anche altri problemi riguardanti tre sistemi fondamentali per l'intelligenza naturale: il sistema immunitario, il sistema comportamentale, il sistema collettivo. Sono notevoli gli sforzi per creare «artificial immune systems» imitando i sistemi protettivi elaborati dall'evoluzione biologica.³⁶ I problemi più ardui però riguardano gli altri due sistemi, strettamente connessi a programmi di forte integrazione tra il programma di un computer e la costruzione di un robot. L'integrazione che nell'intelligenza naturale c'è tra organizzazione neurale e comportamento è molto intensa, decisamente legata alla rilevanza del corpo: «in nature the nervous system evolves simultaneously with the rest of the organism».³⁷ Questo alto livello di integrazione non può essere sottovalutato e sebbene vi siano tentativi sempre più avanzati per realizzare un'integrazione simile nelle macchine, non si può negare che, almeno per ora, rimangono ancora molti interrogativi aperti. Una situazione simile si ha riguardo ai sistemi collettivi, ossia alla rilevanza che per l'evoluzione

ha l'intreccio tra più organismi. Nella dinamica tra competizione e cooperazione, sono emersi modelli di primaria importanza per le dinamiche dei processi adattativi. Il principale compito della tecnologia è di trovare algoritmi ispirati a quei modelli.³⁸ Tra gli aspetti più rilevanti è il doppio tipo di cooperazione intesa come aiuto erogato agli altri: vi sono, infatti, situazioni nelle quali l'aiuto all'altro non implica nessuno costo per il proprio benessere, e situazioni nelle quali l'aiuto all'altro implica dei costi per il proprio benessere.³⁹ Una questione quanto mai rilevante per il nascere e lo svilupparsi di comportamenti morali e che apre un capitolo quanto mai rilevante per il futuro dell'intelligenza artificiale.

Il ruolo del corpo, dei comportamenti e degli scambi rappresentano, a mio avviso, le questioni più rilevanti nel confronto tra intelligenza naturale e intelligenza artificiale. Come si è accennato nel primo paragrafo, si sa l'importanza di quegli aspetti, a cui occorre aggiungere la sfera emotiva. Corpo, azione, emozione, intersoggettività, sono componenti primarie di quella evoluzione biologica che ha reso possibile l'incrementarsi della vita e delle funzioni cognitive. Esistono macchine il cui programma tiene presente quelle componenti e il loro intrecciarsi in un unico complesso sistema dinamico? Secondo alcuni macchine così esistono e tendono anzi a evolversi verso orizzonti per noi quasi impensabili: orizzonti che potrebbero portare alla superintelli-

genza. Con «superintelligenza» si intende indicare l'affermarsi di strumenti tecnologico-digitali capaci non solo di modificare il cervello umano, ma anche di superarlo in modo radicale su molteplici fronti se non addirittura su tutti. La questione è molto discussa e comunque sostenuta soprattutto dalla cosiddetta teoria della singolarità. Alla base della teoria della «singolarità», così com'è stata formulata da R. Kurzweil, sta la convinzione che il progresso tecnologico non sia lineare ma esponenziale, con la conseguenza che a un certo punto, non lontano nel futuro, sarà così esplosivo da modificare tutto ciò che è stato vissuto e pensato nel passato. Il punto di non ritorno sarà costituito dal superamento dei limiti biologici che per tanto tempo hanno caratterizzato l'uomo. Da tener presente che non si tratterà del venir meno della dimensione biologica ma «della fusione fra il nostro pensiero e la nostra esistenza biologica con la nostra tecnologia».⁴⁰ L'assunzione principale della teoria della singolarità è che il nostro universo si basi sull'informazione e che, come vuole il principio antropico,⁴¹ sia segnato da un progetto che ha portato a quel livello di organizzazione dell'informazione che è il cervello umano. Così, se la prima epoca della storia dell'universo è segnata dalla fisica e dalla chimica, la seconda epoca dalla biologia grazie all'avvento della molecola del DNA, la terza epoca dalla comparsa dei neuroni e dalla loro organizzazione in ciò che chiamiamo cervello, la quarta epoca è sempre

più segnata dalla tecnologia e dall'elaborazione dell'intelligenza artificiale. Seguirà la quinta epoca, ossia quella della fusione tra tecnologia e intelligenza umana, tra intelligenza artificiale e intelligenza naturale, e infine la sesta epoca che vedrà l'espandersi dell'intelligenza (umano-tecnologica) su tutto l'universo.⁴² La singolarità ha i suoi presupposti nella svolta tecnologica della quarta epoca e si realizzerà nella quinta e nella sesta epoca.

Il punto nodale della singolarità, come di altre prospettive, è costituito dal programma inteso come organizzazione dell'informazione. L'intelligenza è un programma tanto nella versione biologica quanto nella versione tecnologica. La differenza delle due versioni è dato dal cambiamento del supporto fisico, cerebrale e più ampiamente corporeo. Genetica, nanotecnologia e robotica, «trasformeranno i nostri fragili corpi umani versione 1.0 nei loro corrispondenti versione 2.0, più duraturi e più capaci. Miliardi di nanorobot viaggeranno nella circolazione sanguigna dei nostri organismi e dei nostri cervelli».⁴³ Il programma sopravvive ai suoi supporti fisici nel senso che li trasforma o addirittura li abbandona per nuovi supporti. Il passaggio dall'intelligenza naturale all'intelligenza artificiale non presenta ostacoli perché per la seconda ci si può avvalere dei medesimi meccanismi su cui si basa la prima e che ha portata la prima ai traguardi che osserviamo nella civiltà umana; e grazie alle elaborazioni computazionali di cui sarà capace la macchina

sarà possibile la singolarità, ossia il superamento dei limiti presenti nelle elaborazioni umane. È importante tener presente, però, che tutto si regge sul ciò che viene assunto anche da altri studiosi, e in primo luogo da Minsky, ossia che i sistemi complessi originano da meccanismi relativamente semplici: «Certi fenomeni naturali che possono apparire complessi - scrive Kurzweil - a un qualche livello sono semplicemente il risultato di semplici meccanismi computazionali che essenzialmente sono automi cellulari».⁴⁴ E così, sebbene il cervello umano sia estremamente complesso, si basa su meccanismi più semplici che potrebbero essere individuati e utilizzati per programmare l'intelligenza artificiale, ottenendo così macchine in grado di raggiungere le capacità umane e anche superarle.⁴⁵ In altri termini l'intelligenza, pur essendo emergenziale, ossia sia una realtà più complessa delle sue componenti, abbia una base solida che la costituisce (gli schemi della programmazione) e che le consente di riprodursi in supporti diversi, in corpi diversi, e attraverso questi evolvere in modo esponenziale verso situazioni assolutamente inedite per i corpi biologici. Una prospettiva che incide profondamente anche sul modo di affrontare il fenomeno della coscienza.

2.3. Il programma informatico come base della coscienza nell'uomo e nella macchina

La tesi della singolarità, così come viene esposta da Kurzweil, sembra segnata da una forte marcatura ri-

duzionista e comunque antiolistica. Se l'olismo implica che il tutto sia più della somma delle parti, un atteggiamento esplicitamente antiolistico era già riscontrabile in Minsky quando escludeva questa posizione, e si muoveva verso un evidente riduzionismo nella comprensione della coscienza e delle emozioni.⁴⁶ Ma che ne è, appunto, della coscienza e delle emozioni nella teoria della singolarità? «Le macchine future - si chiede Kurzweil - saranno in grado di avere esperienze emotive e spirituali».⁴⁷ Non è facile rispondere alla domanda, perché ci si trova a confronto con quelle esperienze soggettive che è difficile comprendere anche per le scienze che si occupano del cervello umano. La scienza procede in modo oggettivo, in terza persona, per cui è epistemologicamente difficile affrontare le esperienze soggettive, che sono in prima persona. D'altra parte le esperienze profonde, intime e soggettive del proprio io sono fondamentali e non si possono trascurare, con la conseguenza che macchine veramente evolute dovrebbero possederle. E infatti, secondo il nostro e altri autori, le macchine evolute potranno averle così come le hanno gli esseri umani grazie a quei principi di base dell'intelligenza naturale che sono, o meglio saranno operativi anche nell'intelligenza artificiale. Per comprendere quanto appena detto occorre interrogarsi su «chi sono io», dato che io sono il portatore di quelle esperienze soggettive. Non sono semplicemente l'insieme delle

cellule del mio corpo, dato esse che continuano a mutare mentre io percepisco di essere sempre la stessa persona. Allora il mio io permanente non è dato dalla sua composizione fisica ma dallo schema che opera in essa e attraverso essa: «dato che continuo a cambiare - scrive Kurzweil - sono soltanto uno schema».⁴⁸ Infatti «sono un insieme di materia completamente diverso da quello che ero un mese fa, e tutto quello che rimane è lo schema di organizzazione di quella materia. Anche lo schema cambia, ma lentamente e con continuità».⁴⁹ Le macchine potranno usufruire del medesimo schema anche se su supporti fisici molto diversi rispetto al corpo umano e al cervello biologico. La possibilità, poi, del passaggio dall'uomo alle macchine, che per la precisione sarà una fusione tra biologico e tecnologico, dipende dalla natura dello schema che implica il fenomeno dell'emergenza. Con *emergency* si possono intendere diverse cose.⁵⁰ Per Kurzweil i poteri emergenti riguardano lo schema (il software) e hanno reso possibile quel processo evolutivo che consente allo schema stesso di andare oltre le condizioni biologiche originarie (l'hardware originario). Kurzweil precisa che l'«andare oltre» è ciò che chiamiamo «trascendenza» e che quindi i fenomeni «emergenti» sono la modalità biologica della trascendenza, senza dover ricorrere a dualismi che ipotizzano una trascendenza spirituale al di là di questo mondo. E sono proprio i poteri emergenti dello

schema che consentono un ulteriore passo evolutivo, ossia quello della singolarità in cui si realizzerà il trionfo dell'intelligenza artificiale.⁵¹

La proposta di Kurzweil è antidualistica ma solo in parte. Viene rigettato il dualismo classico tra mondo materiale e mondo spirituale, ma nel confronto tra intelligenza naturale e intelligenza artificiale emergono nuovi dualismi, e precisamente il dualismo tra mente e cervello o tra software e hardware. Il neuroscienziato A. Bell ha fatto osservare che risulta piuttosto problematico il confronto tra il cervello e il computer sottolineando che «a computer is an intrinsically dualistic entity, with its physical set-up designed not interfere with its logical set-up, which executes the computation. In empirical investigation, we find that the brain is not dualistic entity. Computer and program may be two, but mind and brain are one».⁵² La configurazione fisica del cervello interferisce con le funzioni mentali e ne è parte integrante. A questa obiezione, Kurzweil risponde che il dualismo tra struttura fisica e programma logico non è un limite ma il vantaggio della macchina rispetto all'uomo. Egli scrive: «La possibilità di separare in un calcolatore il programma dalla realizzazione fisica che esegue l'elaborazione è un vantaggio, non una limitazione».⁵³ Tutto viene ridotto al vantaggio di avere un software che può essere trasferito a diversi hardware in continuo miglioramento, e che porterebbe anche a riconsiderare la possibilità di trasferire la

stessa mente umana da supporti biologici a supporti non biologici.

Il punto problematico di questa tesi traspare dal testo appena citato ossia nell'intendere il supporto concreto come la semplice «realizzazione fisica» del programma. L'indipendenza del programma dalla realizzazione fisica ha in sé lo strano sapore cartesiano di una *res cogitans* indipendente dalla *res extensa*. Valgono qui le obiezioni di diversi neuroscienziati (è appena il caso di ricordare *L'errore di Cartesio*, di A. Damasio) e in particolare l'obiezione nei confronti di facili paragoni tra intelligenza naturale e intelligenza artificiale. Kurzweil ovviamente insiste sul fatto che «ci sono parallelismi fra cervelli e calcolatori, più di quel che possa sembrare a prima vista»,⁵⁴ ma non pochi studiosi mettono in guardia da esasperare questi parallelismi. «La metafora - scrive F. De Waal - che affianca il cervello a un computer è profondamente fuorviante se pensiamo che il cervello ha milioni di connessioni diverse con il corpo e ne è parte integrante».⁵⁵ Alla base di questo affiancamento sembra stare ancora la vecchia idea di un cervello autosufficiente e padrone del corpo, trascurando altri fattori decisivi: «Celebriamo il cerebrale, crediamo in qualcosa di simile alla "ragione pura" e abbiamo una scarsa opinione delle emozioni, del corpo e di qualunque altra specie che non sia la nostra»,⁵⁶ precisa ancora il primatologo F. De Waal. Alla critica a Cartesio si aggiunge la critica a Kant. Ed in effetti l'idea di un programma indipendente dalla sua

realizzazione fisica ha il sapore di una «ragione pura», ossia di una ragione indipendente dalla realtà fisica, corporea da cui proviene e grazia a cui matura ed evolve. Un'idea che funziona apparentemente bene quando si osservano programma mentale e supporto fisico, mente e corpo, in modo sincronico, trascurando che nel lungo percorso dell'evoluzione il programma non è stato e non è altro che un modo di essere dell'organismo nel suo complesso. Noi osserviamo solo il punto di arrivo e così molti hanno invertito le parti e sono giunti e ancora giungono a ipotizzare un programma che verrebbe prima (una ragione pura) del supporto organico. Un atteggiamento curioso e che si offre a sorprese quando si confrontano l'intelligenza artificiale, per la quale la realizzazione fisica viene effettivamente dopo il programma, e l'intelligenza naturale, per la quale il programma viene dopo il realizzarsi di condizioni fisiche. Una sorpresa espressa dallo stesso Kurzweil quando osserva che mentre nell'evoluzione biologica e nello sviluppo di un individuo compaiono prima le abilità concrete (come la capacità per esempio di allacciarsi le scarpe, di scherzare, o quel buon senso che è riscontrabile già in un bambino) e successivamente le competenze più sofisticate del calcolo matematico, nei computer avviene l'inverso: già i primi computer hanno manifestato le loro abilità di calcolo ma solo recentemente e con grande fatica si sta cercando di costruire macchine (computer-robot)

dotate delle abilità concrete. «Curiosamente - conclude stupito Kurweil - l'evoluzione dell'intelligenza dei calcolatori ha proceduto in direzione opposta a quella della maturazione umana». ⁵⁷ L'inversione non è secondaria quando ci si interroga sulle possibilità di una macchina di essere cosciente, ma Kurzweil la sottovaluta. Confrontandosi con diverse definizioni di coscienza, egli sostiene che «la coscienza è una proprietà emergente di un sistema fisico complesso». ⁵⁸ Ricorrendo alla nozione di proprietà emergente già considerata sopra, secondo cui alla base vi sono principi e regole, già operanti nei sistemi semplici, che rendono possibili i sistemi complessi, Kurzweil sostiene che applicando alla macchina quei principi e quelle regole che nell'uomo hanno portato alla coscienza, consentiranno anche alla macchina di essere cosciente. Altri autori sostengono invece che i fenomeni emergenti implicano indubbiamente una continuità tra i diversi livelli di complessità dei sistemi ma che per i sistemi più complessi occorre considerare alcuni principi e regole non riscontrabili nei sistemi meno complessi.

3. La coscienza come complessità nell'intelligenza naturale e artificiale

Il tema delle complessità non solo è visto con sospetto da chi vuole procedere a un rigido riduzionismo, ma è diversamente inteso anche da chi vi fa riferimento ed è più cauto rispetto a un riduzionismo estremo. Indub-

biamente sotto i sistemi complessi, vi sono elementi che si comportano secondo regole e principi relativamente semplici, ed è un indiscutibile merito decodificare quelle regole e quei principi per spiegare la realtà ed eventualmente riprodurla con tecnologie che si ispirano a quelle stesse regole e a quegli stessi principi. Credo, però, sia opportuno verificare con attenzione fino a che punto tutto fili liscio dal basso all'altro, ossia dal più semplice al più complesso, dato che non sembra che quest'ultimo si lasci addomesticare così facilmente. È il gioco di Archimede: datemi un punto di appoggio e vi sposterò il mondo; ma dov'è quel punto di appoggio se non nel mondo, o nella testa umana che comunque appartiene al mondo, e al mondo nella sua complessità? E qui sorge la tentazione, antica e nuova, di fare del pensiero (la testa umana) un forestiero che viene dall'esterno del mondo per studiare il mondo. Anche lo studio della coscienza viene inghiottito da quella tentazione. E così nello stesso momento in cui si rifiuta ogni spiegazione metafisica della mente umana e della coscienza si ritiene di avere un punto di appoggio che prescinda troppo facilmente dalla rilevanza della sua collocazione nel mondo fisico. Ma cerchiamo di approfondire la questione.

3.1. La coscienza come complessità dell'intelligenza naturale e artificiale

Il tema della complessità e della sua possibile riduzione a regole e principi semplici programmati biologicamente

e riprogrammabili tecnologicamente, rimane una delle più importanti questioni aperte. Da come si è osservato sopra appare evidente che sono in gioco, da una parte la vita che durante l'evoluzione si da delle regole e le regole che, creando programmi, modificano la vita. L'intreccio tra l'evoluzione e il programma costituisce quindi la questione fondamentale. Questione affrontata da diversi studiosi, e recentemente da M. Tegmark. L'evoluzione è il fenomeno in cui il cambiamento avviene secondo paradigmi selettivi (soprattutto la selezione naturale) che portano anche alle capacità cognitive ma che sono indipendenti dall'intervento di un'intelligenza; il progetto è il fenomeno in cui il cambiamento avviene secondo processi tecnologici messi in atto da un'intelligenza. Gli aspetti in gioco sono: a) nell'evoluzione biologica dell'intelligenza naturale il cervello e la mente, ma anche il corpo come contesto originario della relazione del cervello-mente con l'ambiente; b) nel programma tecnologico dell'intelligenza artificiale l'hardware e il software, ma anche il robot come strumento indispensabile nella relazione dell'hardware-software con l'ambiente. Tegmark, come altri, ritiene anzitutto che in entrambi i casi si possa parlare di vita, sia essa naturale o artificiale, biologica o tecnologica, e riconduce tutta la storia della vita alla dinamica tra il supporto fisico (cervello o hardware) e il programma cognitivo (mente o software). Avremmo così anzitutto Vita 1.0, corrispondente allo stadio biologico, dove tanto l'har-

dware quanto il software derivano dall'evoluzione e non da un progetto: durante l'evoluzione la base fisica di ciò che chiamiamo mente, ossia il cervello (hardware) è la condizione necessaria perché evolva anche la stessa mente (software). Avremmo, poi, Vita 2.0, corrispondente allo stadio culturale, dove l'hardware è frutto dell'evoluzione mentre il software è quasi integralmente progettato: grazie al cervello così com'è evoluto, ossia all'hardware, l'uomo può assimilare informazioni ed elaborare algoritmi che incrementano le sue capacità mentali, ossia può progettare il proprio software. Infine, avremmo Vita 3.0, corrispondente allo stadio tecnologico, dove tanto l'hardware quanto il software sono progettati: l'uomo può intervenire tanto sul proprio cervello, anche grazie alla creazione di supporti fisici come il computer, quanto sulle proprie capacità mentali, ossia può progettare tanto il proprio hardware quanto il proprio software.⁵⁹ In altre parole, «Vita 3.0 è padrona del proprio destino, finalmente del tutto libera dai vincoli della sua evoluzione».⁶⁰ Alla base sta la convinzione che l'intelligenza sia un sistema di informazioni capace «di realizzare fini complessi».⁶¹ Per comprendere meglio questa complessità dei fini, si può partire dalla distinzione tra il mondo fisico inorganico e quello organico. Le particelle elementari, di cui è composto l'universo, tendono alla dissipazione, mentre le componenti fondamentali della vita, ossia i geni, tendono alla replicazione, e quindi, almeno entro certi limiti, tendono a

evitare la dissipazione.⁶² Nel passaggio dal mondo inorganico al mondo organico si ha un cambiamento (parziale) dei fini. Quando, però, si giunge a livelli ancora più complessi, come quello del cervello umano, si nota un nuovo cambiamento, dato che l'essere umano si muove in ordine a fini come l'arte, la religione, la scienza, e può anche decidere di non procreare, ossia di sospendere la replicazione. «Il nostro cervello è molto più intelligente dei nostri geni e, ora che comprendiamo il fine dei nostri geni (la replicazione), lo troviamo molto banale e ci risulta facile ignorarlo».⁶³ Il ruolo delle macchine rafforza questo nuovo orientamento dato che hanno il compito di favorire gli scopi umani.⁶⁴ Lo spostamento dei fini nelle forme di intelligenza più avanzata, implica ciò a cui già altri hanno fatto riferimento, ossia l'*emergency*. Il fenomeno forse più emergente del cervello è quello della coscienza che ha indubbiamente proprietà «al di là e al di sopra di quelle delle sue particelle».⁶⁵ Per l'approfondimento di questo aspetto, Tegmark ricorre a una delle tesi neurologiche più recenti sulla coscienza, ossia quella di G. Tononi, che ha lavorato a lungo con G. Edelman. Il cervello è composto di particelle che compongono le molecole e quindi le cellule e i neuroni, ma né le particelle, né le molecole, né le cellule e neppure i neuroni sono coscienti. Solo la complessa organizzazione del cervello, e del cervello in un corpo, mostra la caratteristica della coscienza. La semplice somma di

tutte le caratteristiche delle cellule, sottolinea Tononi, non produce la coscienza perché questa implica una caratteristica aggiuntiva, con la conseguenza che il tutto è più della somma delle parti.⁶⁶ In questo modo viene abbandonata, o almeno così sembra, l'idea di Minsky. Tononi ricorre alla tesi secondo cui il tutto è più della somma delle parti attraverso la teoria dell'*informazione integrata* secondo cui l'informazione è generata, appunto, da un sistema oltre le sue parti.⁶⁷ Il suo simbolo può essere Φ (Phi), ossia il simbolo del rapporto aureo: il modo giusto per dividere qualcosa in parti. In termini più dettagliati: a) ricorda la fenomenologia; b) la I rimanda all'idea di informazione; c) O, il cerchio, richiama l'integrazione.⁶⁸

La coscienza ha a che fare con l'informazione integrata perché implica «un sistema per il quale l'informazione generata dal tutto al di là delle sue parti raggiunge il massimo [...]. Un complesso».⁶⁹ Non si può parlare di coscienza senza riferirsi a un sistema abbinato all'idea di *complesso*: «un complesso è dove vive la coscienza».⁷⁰

La comprensione della coscienza secondo i criteri della complessità e quindi come fenomeno emergente, è quanto mai interessante ma è comunque sempre inteso entro il quadro di riferimento che concepisce l'intelligenza come informazione. Ciò porta Tegmark a intendere l'emergenza come (relativa) indipendenza dell'intelligenza rispetto al suo supporto fisico. Di conseguenza anche la coscienza è indipendente dal suo sup-

porto fisico.⁷¹ Per la coscienza importa solo la struttura dell'elaborazione dell'informazione, non la struttura della materia che compie l'elaborazione. E qui torna la nozione di emergenza. Tegmark infatti sostiene che «*se la stessa elaborazione dell'informazione obbedisce a certi principi, può dar luogo al fenomeno emergente di livello superiore che chiamiamo coscienza*».⁷²

Ma come può avvenire questo fenomeno emergente di livello superiore? Su questo punto Tegmark ritiene che «il sistema cosciente debba essere integrato in un tutto unificato, come ha sostenuto Giulio Tononi, perché, se fosse costituito da due parti indipendenti, queste si sentirebbero come due entità coscienti separate».⁷³ Tegmark sostiene che alla base della coscienza ci siano almeno quattro principi (di informazione, di dinamica, di indipendenza, di integrazione) e, sulla base del tutto irriducibile alle sue parti, sostiene che «i quattro principi insieme significano che un sistema è autonomo, ma le sue parti non lo sono».⁷⁴

La teoria di Tononi, dell'informazione integrata, è stata molto criticata, ma mantiene un certo valore dato che tutte le altre teorie sulla coscienza sono decisamente vaghe. Uno degli aspetti più interessanti ma anche problematici è che Tononi crede sia possibile estendere la coscienza all'intelligenza artificiale a condizione che si mettano insieme tutti quegli elementi grazie ai quali scatta il «di più» di quel tutto (rispetto alla somma delle sue parti) che costituisce un sistema cosciente. Infatti se quel «di

più» scaturisce automaticamente dall'assemblare tecnologicamente (informaticamente) tutti gli elementi che lo rendono possibile nell'ambito biologico, lo si può ancora considerare veramente un «di più» rispetto alla somma delle parti? Naturalmente occorre verificare meglio di quali elementi si tratti, ma sembra strano, però, che siano riducibili a quelli assemblabili in un programma che predeterminerebbe la coscienza.

3.2. La coscienza come esperienza complessa del corpo

In un prestigioso manuale di diversi anni fa leggiamo: «Gli studi sull'intelligenza artificiale e le analisi sui meccanismi di riconoscimento delle forme impiegate dai computer hanno permesso di arrivare alla conclusione che il sistema nervoso riconosce il movimento e le forme impiegando sistemi di analisi che nessun computer oggi esistente è in grado di simulare. La semplice osservazione del mondo esterno, il riconoscimento di un viso o di un paesaggio comportano una capacità di analisi molto maggiore di quella necessaria per risolvere problemi di logica o per giocare a scacchi». ⁷⁵ I progressi più recenti sono considerevoli ma sembrano confermare che il «riconoscimento di un viso» implichi comunque una capacità di analisi molto maggiore di quanto occorre per risolvere problemi logici. Ciò non toglie che un computer decisamente più abile dell'uomo a risolvere problemi logici, possa riconoscere un viso e che possa rag-

giungere e forse superare l'uomo in tutti gli aspetti dell'attività cognitiva, compresa la coscienza. Si può, però, sollevare qualche dubbio se il raggiungimento di questi traguardi, e soprattutto quello di una macchina cosciente, possa trascurare la rilevanza di alcune componenti così rilevanti per l'uomo cosciente. In modo particolare la rilevanza del corpo, delle emozioni e dell'intersoggettività. La rilevanza del corpo per la conoscenza e soprattutto per la coscienza nella sua valenza di esperienza soggettiva è stata richiamata da diversi autori appartenenti all'ambito della filosofia e della scienza, e in modo particolare della fenomenologia e delle neuroscienze. Già da tempo e con una sensibilità di tipo fenomenologico è stato fatto osservare che la capacità del cervello di apprendere viene dal suo inserimento in un corpo che gli fa da interfaccia con l'ambiente. ⁷⁶ Questa osservazione, portata avanti anche da alcuni esponenti delle neuroscienze, e in parte già segnalata sopra, solleva seri dubbi sulla possibilità di avere delle macchine che pur superando di gran lunga le capacità di calcolo dell'essere umano ne uguagliano altre capacità cognitive e in particolare la coscienza, almeno nella misura in cui gli specialisti nell'intelligenza artificiale sottovalutino ciò che negli esseri umani è il corpo. Il punto decisivo è che non si può limitare la questione dell'intelligenza al rapporto tra le funzioni cognitive e l'organo che costituisce il loro supporto fisico: in poche parole non si può limitare l'intelligenza al rapporto

tra la mente e il cervello. In questo rapporto, infatti, entra in gioco una terza componente, che costituisce l'entità in cui risiedono il cervello e quindi la mente, ossia il corpo. Si è visto che al cervello e alla mente dell'intelligenza naturale sono stati fatti corrispondere l'hardware e il software dell'intelligenza artificiale. Quando ci si è accorti della necessità, per le macchine intelligenti di mettere in gioco anche il robot, si è in qualche modo riconosciuto quanto sia rilevante che, almeno a prima vista, corrisponde al corpo biologico. Nell'ambito dell'intelligenza artificiale è apparso, infatti, sempre più chiaro che non è sufficiente giocare solo le carte del computer, ossia l'hardware e il software, ma che ne occorre una terza, il robot. Una macchina che si muove e impatta l'ambiente è una fonte fondamentale per lo sviluppo cognitivo del computer, così come un corpo che si muove nell'ambiente è stato ed è fondamentale per lo sviluppo delle capacità cognitive dell'uomo.

Un aspetto fondamentale è che quando si parla della rilevanza del corpo per lo sviluppo dell'intelligenza naturale ci si riferisce anche all'impatto emotivo con l'ambiente biologico e con l'ambiente sociale. Di conseguenza ci si riferisce anche alla dinamica intersoggettiva. Le capacità cognitive dell'uomo e soprattutto quella sua esperienza soggettiva che chiamiamo coscienza sono strettamente legate al corpo per la sua rilevanza emotiva e intersoggettiva.

Per questo alcuni esponenti della ricerca sull'intelligenza artificiale hanno insistito perché i colleghi tengano conto delle emozioni e della socialità.⁷⁷ Il rapporto della mente e del cervello col corpo e le sue dinamiche emotive e intersoggettive è, però, molto più complesso di quanto sembrano ritenere alcuni informatici meno avvertiti. Ed è proprio su ciò che si intende per complessità che si gioca il confronto più importante.

Nell'ambito dell'intelligenza naturale la complessità è l'indistricabile rapporto del cervello e della mente col corpo, anzitutto perché il cervello è «corpo nel corpo». Con ciò intendo dire che nel caso dell'intelligenza naturale è il corpo che si auto-organizza in modo da avere un cervello e di farlo evolvere come mente, ossia come capacità cognitiva e anche come coscienza: il corpo, nel suo impatto con l'ambiente, fa il cervello come organo del pensiero, delle emozioni, dei sentimenti, della coscienza. Sotto questo profilo la mente e non solo il cervello è corpo nel corpo. Il cervello è corpo come organo, la mente è corpo come funzione: ma entrambi sono corpo, e più precisamente corpo nel corpo. Questo intreccio profondo rende possibile anche che il corpo sia mappato dal cervello e che questa mappatura sia già mente, per cui a partire dal fatto che il cervello e la mente sono nel corpo, si sviluppa il fenomeno per cui il corpo è nel cervello e nella mente.

Nel caso dell'intelligenza artificiale si parte da un hardware e da un software

che precedono, per così dire il loro corpo esterno, ossia il robot. Indubbiamente, nel caso dell'uomo, se è anzitutto l'intero corpo che rende possibile il cervello e la mente (il cervello e la mente nel corpo), queste, a loro volta, svolgono il compito di dirigere il corpo (il corpo nel cervello e nella mente). Anzi dovremmo dire, il resto del corpo, dato che il cervello e la mente sono parte del corpo. È come se il corpo, tutto il corpo (ossia nella sua complessità) si servisse del cervello e della mente per dirigersi. Se è così facile cadere nell'ingenuità di parlare di un cervello che dirige il corpo, quasi postulando un cervello che dirige il corpo dal di fuori, dimenticando che a dirigere è comunque una parte del corpo, probabilmente è perché il cervello ci rende coscienti del corpo ma non del cervello stesso. Il cervello sa del corpo ma non di se stesso come luogo del sapere, e così grazie al cervello siamo coscienti di avere un corpo ma non di avere un cervello che pensa (sarà la scienza a scoprirlo). Il cervello si nasconde e così favorisce una sorta di dualismo tra sé e il resto del corpo, finché la ricerca non mette bene in evidenza che il cervello è l'organo indispensabile per la mente, e che lo è grazie ai processi di evoluzione e di sviluppo attraverso cui passa il corpo di cui fa parte. Occorre, quindi, evitare ogni dualismo, e ribadire che è il corpo che si serve del cervello (e della mente) per dirigere se stesso. La dinamica è corpo → cervello come parte del corpo → corpo.

L'intelligenza artificiale, invece, quando intreccia computer e robot, parte subito dal computer i cui programmi dirigono il robot: computer (= cervello) → robot (= corpo). È infatti anzitutto il computer che dirige il robot anche se successivamente il robot influenza il programma del computer. Questo deve far ripensare Vita 1.0, 2.0, 3.0, e il binomio evoluzione-progetto non solo entro l'orizzonte hardware-software, corrispettivi di cervello-mente, ma anche entro l'orizzonte del trinomio corpo-cervello-mente, dove il corpo è allo stesso tempo evoluzione e progetto. Il fatto che con lo sviluppo della tecnologia si intervenga sempre più marcatamente sul corpo, ossia che il cervello-mente modificano il corpo, non può trascurare che tutto è partito dall'appartenenza del cervello-mente al corpo e che questa appartenenza non è occasionale ma intrinseca alle capacità umane sia pure in modalità diverse, come diversa è la capacità oggettiva di calcolo e l'esperienza soggettiva della coscienza.

L'approfondimento della questione richiede un confronto con quello che è il punto chiave dell'intelligenza artificiale, rilevabile anche nell'intelligenza naturale: ossia il programma come organizzazione di informazioni. Un sistema intelligente artificiale è un programma e quindi presuppone un programmatore. La prima considerazione che viene in mente se ci si avvia verso un confronto con l'intelligenza naturale è che quest'ultima non è programmata da un'intelligenza antecedente. All'origine di tutto non

c'è un programma ma un'evoluzione piuttosto casuale tracciata dalla selezione naturale e da altre selezioni. La selezione non è il programma elaborato da un'intelligenza, ma un processo che di fatto, almeno sulla Terra, ha portato alla comparsa dei neuroni e col tempo all'elaborazione di quell'intelligenza piuttosto sofisticata che si può osservare in un organismo dotato di un cervello. L'aspetto interessante è che è stata l'intelligenza non programmata, l'intelligenza naturale, a programmare l'intelligenza programmata, l'intelligenza artificiale, con la probabile conseguenza che quanto è più pertinente ai processi di programmazione, come il calcolo logico, possa trasferirsi senza problemi dall'uomo alla macchina, mentre altre capacità in cui diminuisce o scompare la rilevanza del programma rendono piuttosto problematico quel trasferimento. Il programma, inteso come organizzazione che gestisce e controlla l'informazione in modo oggettivo, diventa proprio ciò che non può rendere conto di fenomeni che lo anticipano, come la causalità e le non controllabilità dell'esperienza soggettiva quale la coscienza.

In questa problematica rientra proprio il concetto di singolarità ma in termini molto diversi da quelli di Kurzweil. Per questo autore, la singolarità è il punto di svolta storicamente unico provocato dalla fusione tra intelligenza naturale e intelligenza artificiale sulla base della possibilità di trasferire un programma (diciamo la mente) da un supporto all'altro (diciamo da un

corpo all'altro). Ma forse, quella fusione, o comunque il confronto tra l'uomo e la macchina dovrebbe tenere presente un'altra singolarità: quella di Carlo, di Diana, di Luisa. Si tratta di un punto decisivo perché alla base dell'interscambio tra uomo e macchina sta la convinzione che ciò avvenga sulla base di processi computazionali universali. Per quanto riguarda il computer, si osserva che sebbene possa raggiungere sviluppi ben superiori ai suoi inizi si basi comunque sui medesimi processi computazionali di fondo degli inizi (tanto per intenderci quelli di Turing). Lo scambio e la fusione tra il computer e il cervello umano dipende dal fatto di riconoscere anche in quest'ultimo quei processi computazionali universali.⁷⁸ In questo modo si sottovaluta che i meccanismi operativi a un livello della realtà, come i sistemi informatici e molecolari, sono indubbiamente necessari per gli sviluppi di livelli più complessi della realtà ma non sono in grado di spiegare tutto ciò che avviene a questi livelli più complessi, come la coscienza e l'esperienza soggettiva e singolare. Ma soprattutto sono necessarie ma insufficienti per comprendere in modo rigoroso e completo il fenomeno della differenza delle singole persone, che invece è da ritenersi essenziale per lo sviluppo dell'intelligenza di alto livello e che probabilmente costituisce il vertice dell'evoluzione biologica. L'autodifesa e lo sforzo di conservazione della singola cellula sono prerequisiti della coscienza e della singolarità ma non

sono sufficienti per la piena e rigorosa spiegazione della complessità riscontrabile nell'esperienza di un mammifero, di un primate e soprattutto di quel primate che è l'uomo.

Nulla impedisce che quel vertice e questa complessità siano raggiunte anche dalle macchine, ma anche per loro, probabilmente, deve valere la percezione della propria differenza specifica in modo che ogni singola macchina sia consapevole della propria singolarità e unicità rispetto alle altre macchine. L'evoluzione biologica ha portato l'uomo a questa consapevolezza e questo è descrivibile come un programma (estremamente complesso), ma un programma che non parte da un'intelligenza precedente che l'abbia programmata (a meno che non si voglia mettere in gioco Dio come componente intrinseca alla scienza). L'unicità di ogni singola persona è così strettamente legata al tema della coscienza che una macchina cosciente non può sottovalutarla. E ciò significa che la macchina senta l'esigenza linguistica di elaborare nomi propri irriducibili ai nomi comuni.

NOTE

¹ Giorgio Bonaccorso è monaco benedettino e docente di Antropologia ed Estetica rituale presso l'Istituto di Liturgia Pastorale di Santa Giustina (Padova) e in altri Istituti Teologici. È studioso delle radici profonde del rito, con una particolare attenzione all'aspetto antropologico e al dialogo con le scienze cognitive e le neuroscienze.

² Cf DE WAAL Frans, *Siamo così intelligenti da capire l'intelligenza degli animali?*, Milano, Raffaello Cortina 2016.

³ Cf *ivi* 28-31.

⁴ Se per un verso ogni organismo appartiene a una determinata nicchia ecologica, per un altro verso, come osserva J. von Uexfüll ricorrendo al termine *Umwelt* (mondo circostante, ambiente), «percepisce l'ambiente nel suo modo proprio» (*ivi* 20).

⁵ «Nell'insieme, pare che le differenze neurali siano insufficienti a fare dell'unicità umana una conclusione scontata», *ivi* 161.

⁶ Cf *ivi* 204.

⁷ Il 7 luglio 2012 un gruppo di neuroscienziati cognitivi, neurofarmacologi, neurofisiologi, neuroanatomisti e neuroscienziati computazionali riuniti nell'università di Cambridge, hanno dichiarato che allo stato attuale delle ricerche non vi sono motivi per escludere l'esperienza cosciente negli animali non umani (cf *The Cambridge Declaration on Consciousness*, in <http://worldanimal.net/images/stories/documents/Cambridge-Declaration-on-Consciousness.pdf> [24-10-2020]).

⁸ DE WAAL, *Siamo così intelligenti da capire l'intelligenza degli animali?* 295.

⁹ Cf *Id.*, *L'ultimo abbraccio. Cosa dicono di noi le emozioni degli animali*, Milano, Raffaello Cortina 2020, 210.

¹⁰ Si pensi ai lavori di J. Panksepp, A. R. Damasio, J. LeDoux, M. Jaennerod. Si vedano i numerosi interventi in ARMONY Jorge - WUILLEUMIER Patrik (eds), *The Cambridge Handbook of Human Affective Neuroscience*, New York, Cambridge University Press 2013.

¹¹ DE WAAL, *L'ultimo abbraccio* 109.

¹² *Ivi* 110.

¹³ DE WAAL, *Siamo così intelligenti da capire*

l'intelligenza degli animali? 170.

¹⁴ Cf *ivi* 171. «Nei bambini una comprensione dei bisogni e dei desideri si sviluppa vari anni prima che si rendano conto di ciò che gli altri fanno. Essi leggono i sentimenti degli altri molto tempo prima dei loro pensieri. Questo fatto ci fa pensare che siamo su una strada sbagliata quando tentiamo di esprimere tutto questo in termini di pensiero astratto o di teorie sugli altri» (*ivi* 175).

¹⁵ DE WAAL, *L'ultimo abbraccio* 115.

¹⁶ Cf JEANNEROD Marc, *Motor Cognition: What Actions Tell the Self*, Oxford, Oxford University Press 2006, 72-74.

¹⁷ Vi sono diversi esperimenti sulla rilevanza dell'imitazione e del riconoscimento dello scopo altrui, cf FRITH Chris, *Inventare la mente. Come il cervello crea la nostra vita mentale*, Milano, Raffaello Cortina 2009, 183-188.

¹⁸ Su questo campo sono state centrali le ricerche di LIBET Benjamin, *Mind Time. Il fattore temporale nella coscienza*, Milano, Raffaello Cortina 2007.

¹⁹ FRITH, *Inventare la mente* 236.

²⁰ *Ivi* 236-237.

²¹ *Ivi* 237.

²² *Ivi* 238.

²³ KURZWEIL Ray, *La singolarità è vicina*, Santarcangelo di Romagna, Maggioli 2014, 4.

²⁴ Un'espressione di questa svolta sarebbe la nascita, nel 1975, della sociobiologia con Edward O. WILSON (*Sociobiologia. La nuova sintesi*, Bologna, Zanichelli 1983⁹), cf STORRS Hall J., *Beyond IA: Creating the Conscience of the Machine*, New York, Prometheus Books 2007, 105.

²⁵ GALLINA Paolo, *La mente liquida. Come le macchine condizionano, modificano o potenziano il cervello*, Bari, Dedalo 2019, 28.

²⁶ MINSKY Marvin, *La società della mente*, Milano, Adelphi 1990², 34.

²⁷ Cf *ivi* 42-45.

²⁸ Si veda orientativamente CAWSEY Alison, *The Essence of Artificial Intelligence*, Upper Saddle River, Prentice Hall 1998.

²⁹ Cf HÉNIN Silvio, *Al. Intelligenza artificiale tra incubo e sogno*, Milano, Hoepli 2019, 57.

³⁰ Si veda la questione in KAPLAN Jerry, *Intelligenza artificiale. Guida al futuro prossimo*, Roma, Luiss University Press 2018².

³¹ Si pensi, per esempio, all'ampio lavoro di FLOREANO Dario - MATTIUSI Claudio, *Bio-inspired artificial intelligence. Theories, Methods, and Technologies*, Cambridge - London, The MIT Press 2008.

³² Cf *ivi* 1; 13-14.

³³ Per la questione cf *ivi* 139 e le pagine seguenti.

³⁴ Cf *ivi* 175. In tutto ciò è in gioco l'elaborazione di algoritmi evolutivi (cf *ivi* 238).

³⁵ Cf *ivi* 256. Un aspetto decisivo è costituito dall'evodevo, ossia dal parallelismo tra l'evoluzione filogenetica (*evolution*) e lo sviluppo ontogenetico (*development*): parallelismo osservabile nell'ambito biologico e riproducibile nell'ambito tecnologico (cf *ivi* 298).

³⁶ Cf *ivi* 336.

³⁷ *Ivi* 499.

³⁸ Cf *ivi* 515.

³⁹ Cf *ivi* 549.

⁴⁰ KURZWEIL, *La singolarità è vicina* 9.

⁴¹ Cf BARROW John David, *Teorie del tutto. La ricerca della spiegazione ultima*, Milano, Adelphi 1992; *Id.*, *Il mondo dentro il mondo*, Milano, Adelphi 1992²; *Id.*, *Dall'io al cosmo. Scienza, arte, filosofia*, Milano, Raffaello Cortina 2000.

⁴² Cf KURZWEIL, *La singolarità è vicina* 15-21.

⁴³ *Ivi* 296.

⁴⁴ *Ivi* 85.

⁴⁵ Cf *ivi* 103-196.

⁴⁶ Cf MINSKY, *La società della mente* 100; 315.

⁴⁷ KURZWEIL, *La singolarità è vicina* 375.

⁴⁸ *Ivi* 381.

⁴⁹ *Ivi* 382.

⁵⁰ Per il lungo dibattito sull'emergenza, soprattutto nel suo rapporto dialettico col riduzionismo, cf CHIBBARO Sergio - RONDONI Lamberto - VULPINI Angelo, *Reductionism, Emergence and Levels of Reality. The Importance of Being Borderline*, Heidelberg, Springer 2014.

⁵¹ Cf KURZWEIL, *La singolarità è vicina* 386-387.

⁵² BELL Anthony J., *Levels and loops: the*

future of artificial intelligence and neuroscience, in *Philosophical Transactions B*, 354(1999) 2019.

⁵³ KURZWEIL, *La singolarità è vicina* 447.

⁵⁴ Id., *Come creare una mente. I segreti del pensiero umano*, Santarcangelo di Romagna, Maggioli 2019, 12.

⁵⁵ DE WAAL, *L'ultimo abbraccio* 256.

⁵⁶ *Ivi* 257.

⁵⁷ KURZWEIL, *Come creare una mente* 150.

⁵⁸ *Ivi* 170.

⁵⁹ Cf TEGMARK Max, *Vita 3.0. Essere umani nell'era dell'intelligenza artificiale*, Milano, Raffaello Cortina 2018, 44-50.

⁶⁰ *Ivi* 49.

⁶¹ *Ivi* 61.76.

⁶² Cf *ivi* 322.

⁶³ *Ivi* 325.

⁶⁴ Cf *ivi* 347.

⁶⁵ *Ivi* 377.

⁶⁶ Cf TONONI Giulio, *Phi. Un viaggio dal cervello all'anima*, Torino, Codice 2017, 140.

⁶⁷ Cf *ivi* 146.

⁶⁸ Cf *ivi* 147.

⁶⁹ *Ivi* 150.

⁷⁰ *L. cit.*

⁷¹ Cf TEGMARK, *Vita 3.0*, 381.

⁷² *Ivi* 382.

⁷³ *Ivi* 383.

⁷⁴ *L. cit.*

⁷⁵ KANDEL Eric R. - SCHWARTZ James - JESSELL Thomas M., *Fondamenti delle neuroscienze e del comportamento*, Milano, Ambrosiana 1999, 387.

⁷⁶ Su questo punto ha insistito già tempo fa Dreyfus in modo molto critico (cf DREYFUS Hubert L., *What Computers Still Can't Do. A Critique of Artificial Intelligence*, Cambridge [MA], MIT Press 1992, 235-255).

⁷⁷ Su cui insiste Fei-Fei Li (cf FORD Martin, *Architects of Intelligence: The Truth About AI from the People Building it*, Kindle, Packt Publishing 2018, 147). Il testo è una serie di interviste con numerosi studiosi dell'intelligenza artificiale.

⁷⁸ Cf STORRS, *Beyond IA* 119-131.