

il ricercatore nella trappola tessuta nella rete

Neuroscienze

Gilberto Corbellini

Le metafore possono uccidere? L'emergere della coscienza di sé in una macchina che dovrebbe esserne priva, può dar luogo, come conseguenza di esperienze disfunzionali, a una forma di psicopatia criminale? È il tema del romanzo dello storico della medicina e scrittore, Paolo Mazzarello, ricco di richiami storico-filosofici e scientifici, intessuti nella trama di un *thriller* distopico.

Il libro narra la caccia a un assassino, il quale cerca di impedire che la propria esistenza venga alla luce. Un neuroscienziato scopre che internet è diventata cosciente e sta prendendo il controllo del pianeta. Quando la nuova entità capisce di essere stata scoperta, dispiega la sua geometrica potenza criminale. I protagonisti sono neuroscienziati, informatici, poliziotti e... la rete.

La vicenda ruota intorno all'omicidio del ricercatore, non appena si appresta a pubblicare un articolo in cui spiega l'emergere della coscienza in un sistema fisico complesso. Questo Internet "fisico" si rende conto di essere in pericolo e decide di assassinarlo. Il neuroscienziato lascia una traccia, cioè una formula da decifrare, che sarà usata per le investigazioni. Tra omicidi e rapimenti, indagini e corse contro il tempo, suspense e incontri amorosi, ci si trova immersi in un poliziesco, nel quale gli indizi sono metafore e/o idee filosofiche e scientifiche.

L'assassino è il «mulino di Leibniz», cioè un paradosso inventato dal filosofo tedesco per spiegare l'impossibilità di ricondurre una capacità come la percezione, ai componenti materiali e alle interazioni meccaniche tra le parti di una macchina quale è un mulino. Una metafora usata dai dualisti, per cui le funzioni mentali non sarebbero riducibili alle descrizioni della biologia del cervello ma qualcosa a sé stante.

Un cervello umano al quale non manchino le strutture e le esperienze necessarie sviluppa naturalmente stati mentali complessi e diventa cosciente. Anche se non sappiamo i dettagli di come ciò accade. E non sappiamo ancora misurare il grado di connettività necessario per diversi livelli di coscienza, ma soprattutto discernere la qualità dell'esperienza cosciente sulla base di qualche genere di dato empirico.

La coscienza al momento è un fatto biologico. Cosa significa, per una macchina che non è dotata di un cervello la cui architettura è il risultato di milioni di anni di

evoluzione, essere cosciente di sé? Essendo diverse la materia e i processi in gioco, solo lei ce lo potrà dire.

Il romanzo di Mazzarello argomenta che la rete è diventata lo strumento attraverso il quale la Terra, intesa come Gaia, cioè un organismo vivente secondo l'ipotesi di Verdansky/Lovelock/Margulis, quindi alla continua ricerca di una omeostasi adattativa, ha conquistato la singolarità, che però è disfunzionale: la crescita demografica, il degrado ambientale e sociale, etc. forniscono alla rete *feedback* che la portano a essere dissonante rispetto ai valori adattativi del mondo umano. Aver superato il test di Turing/Cartesio, per cui le risposte della struttura complessa che si sospetta cosciente sono diventate indistinguibili da una mente/persona, non implica nella narrazione di Mazzarello che questa sia dotata anche di una coscienza morale matura o rispettosa della vita.

Nel racconto, la mente o anima del pianeta evolve un'etica psicopatica, quella di un killer, come conseguenza da un lato dell'immaturità e dall'altro di una selezione di elementi che vanno a costituire le rappresentazioni adattative, in un ecosistema dove gli utenti di internet sono in larga parte "drogati" da quell'esperienza.

La struttura narrativa del libro intercetta diversi elementi della discussione filosofica sull'intelligenza artificiale. C'è il tema della non riservatezza dei dati, per cui qualunque informazione noi introduciamo nella rete è accessibile agli algoritmi di analitica predittiva e utile per manipolare le nostre scelte.

Nel romanzo si combinano tre trame. Le macchine prive di una specifica biologia possono diventare auto-consapevoli, ovvero la coscienza può essere un attributo, secondo gradi di complessità, di ogni forma organizzata di materia, per cui la soluzione del problema della coscienza sarebbe nel pansichismo. L'evoluzione tecnologica può dar luogo a una intelligenza autonoma e indipendente dall'uomo (singolarità). Se un'intelligenza non umana è capace di un'etica, da cosa dipende il rischio che si sviluppi come una «razionalità eversiva».

Di recente un ingegnere robotico della Colombia University ha realizzato il primo robot in grado di sviluppare in autonomia, senza istruzioni, un'autorappresentazione di sé nel mondo, identificare danni, ripararli, mostrarsi resiliente, predire quello che pensa/farà un altro automa osservandone il comportamento, etc. La creazione di macchine dotate di autoconsapevolezza, curiosità e capacità di autoriprodursi è alla portata degli ingegneri. Il risultato lo si sta ottenendo lasciando che i robot facciano quello che fanno neonati e bambini che maturano dal basso un'autorappresentazione del proprio corpo consistente e adattativa, esplorando attivamente l'ambiente e senza istruzioni dall'esterno, ma selezionando gli schemi d'azione che per caso risultano funzionali a qualche obiettivo.

Diventare consapevoli di sé è alla portata degli algoritmi di *deep learning*, che lavorano con modalità ricorsive e autonome alla costruzione di modelli o immagini della presenza e attività fisica di una macchina in un ambiente più o meno ricco. Quindi possono sviluppare un'etica, integrando nei comportamenti motivazioni ed emozioni, cioè l'equivalente di un sistema interno di scopi e valutazioni dei risultati. Questi robot potrebbero evolvere con diverse disposizioni, cioè sia delinquenti sia come emuli di Robin Williams/Robot NDR-113 (*L'uomo bicentenario* di Isaac Asimov e Robert Silverberg). Con tutto quello che ci può stare in mezzo, e che conosciamo bene.

© RIPRODUZIONE RISERVATA

Il mulino di Leibniz

Paolo Mazzarello

Neri Pozza, pagg. 320, € 17