

Benjamín Labatut

GLI DÈI DELLA LOGICA

Prima e dopo l'intelligenza artificiale



Benjamín Labatut

GLI DÈI DELLA LOGICA

Prima e dopo l'intelligenza artificiale

Non sapremo mai quanti morirono durante la Jihad butleriana. Erano milioni? Miliardi? Forse trilioni? Fu una collera fantastica, una grande rivolta che si propagò come un incendio, consumando tutto ciò che incontrava sul suo cammino, un caos che travolse intere generazioni in un'orgia di distruzione durata quasi cento anni. Una guerra con un bilancio di vittime così alto da lasciare una cicatrice permanente nell'anima dell'umanità. Ma non conosceremo mai i nomi di quelli vi combatterono e morirono, né le immense sofferenze e la distruzione che provocò, perché la Jihad butleriana, per quanto abominevole e devastante, non c'è mai stata.

La Jihad era un evento immaginario, inventato da Frank Herbert come parte della storia che anima la sua saga di fantascienza, *Dune*. Si trattava dell'ultima resistenza dell'umanità contro la tecnologia senziente, una crociata per rovesciare il dio della logica delle macchine e sradicare i computer e i robot coscienti che nel futuro ci avevano quasi completamente schiavizzati. Herbert la descrisse come «una pausa talamica nella storia dell'umanità», un'epoca di tale violenza incontrollata che ha trasformato completamente il modo in cui la società si è sviluppata da allora in avanti. Ma sappiamo molto poco di ciò che accadde realmente durante la battaglia perché nella serie originale di *Dune*, Herbert ci fornisce solo i contorni più vaghi: accenni, mormorii e sussurri che portano il peso spettrale della profezia. La Jihad ha rimodellato la civiltà mettendo al bando l'intelligenza artificiale o qualsiasi macchina che simulasse le nostre menti, ponendo un freno ai peggiori eccessi della tecnologia. Tuttavia, è stata combattuta così tanti eoni prima degli eventi descritti nel romanzo che a quel tempo era ormai sbiadita nella leggenda e cristallizzata in libri apocrifi. Le lezioni duramente apprese dalla catastrofe sono preservate nella saggezza popolare e in alcuni detti: «L'uomo non può essere sostituito»; «Un tempo gli uomini dedicavano il proprio pensiero alle macchine, nella speranza che esse li avrebbero liberati. Ma questo consentì ad altri uomini di servirsi delle macchine per renderli schiavi»; «Non ci fidiamo dell'ignoto che può nascere dalla tecnologia creativa»; «Dobbiamo annichilire le macchine che pensano». Il lascito più duraturo della Jihad è stato un profondo cambiamento nel rapporto dell'umanità con la tecnologia. Infatti l'obiettivo di questa grande caccia, quando perseguitammo e depredammo quegli stessi artefatti che avevamo creato per poterci sedere al posto che la natura ci aveva assegnato, non era soltanto l'intelligenza meccanica ma l'atteggiamento macchinico che aveva preso possesso della nostra specie: «Gli umani avevano lasciato che le macchine usurpassero il nostro senso della bellezza, la nostra necessaria identità, in base alla quale formiamo i nostri giudizi», scrisse Herbert.

«Gli uomini devono decidere in prima persona la propria vita. Le macchine non sapranno mai far questo. Il ragionamento dipende da chi programma, non dalla ferraglia che esegue ciecamente. E noi siamo il programma supremo!»

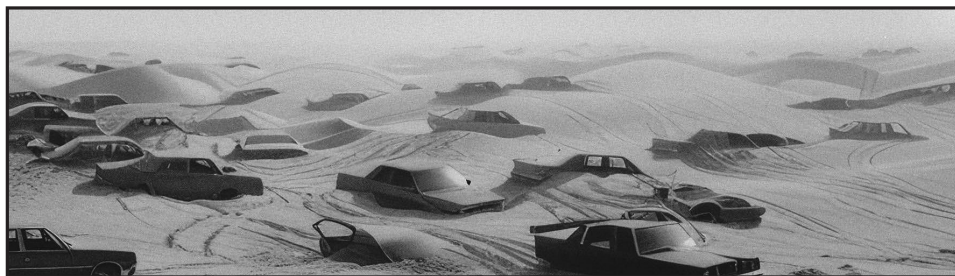
La Jihad butleriana rimosse una stampella – quella parte di noi che avevamo ceduto alla tecnologia – e costrinse le menti degli umani a svilupparsi ben oltre i limiti del ragionamento meccanicistico, in modo da non dover più dipendere dai computer nel farli pensare al posto nostro. La fantasia di Herbert, la sua visione lungimirante di una guerra devastante tra l'umanità e il dio della logica delle macchine, sembrava alquanto bizzarra nel momento in cui cominciò a scrivere la saga, negli anni Sessanta. A quell'epoca i computer erano primitivi secondo gli standard moderni, marchingegni formati da enormi calcolatori che potevano elaborare solamente centinaia di migliaia di cicli al secondo (anziché miliardi di cicli, come oggi), avevano pochissima memoria, funzionavano tramite schede perforate e non erano collegati tra loro. E fin da allora abbiamo tranquillamente ignorato gli avvertimenti di Herbert, ma ora la Jihad butleriana è improvvisamente tornata a tormentarci. L'apocalisse dell'intelligenza artificiale è una nuova paura che tiene molte persone sveglie la notte, un terrore nato dai grandi progressi che sembrano suggerire che, se non siamo molto attenti, potremmo – con le nostre mani – generare un futuro in cui non ci sarà posto per l'umanità. Questo strano incubo è un pericolo credibile dal momento che molti dei nostri sogni minacciano di realizzarsi. È il culmine di un lungo processo che risale alle origini della civiltà stessa, all'epoca in cui il mondo era imbevuto di magia e di terrore, e l'unico modo per garantire la nostra sopravvivenza era quello di invocare il potere degli dèi.

L'apoteosi ha sempre perseguitato l'anima degli esseri umani. Fin da tempi remoti abbiamo patito la brama di diventare come gli dèi e di superare i limiti che la natura ci ha imposto. Per raggiungere questo obiettivo, abbiamo costruito altari e compiuto rituali per invocare saggezza, benedizioni e i mezzi per oltrepassare le nostre capacità. Anche se tendiamo a credere che soltanto adesso, nel mondo moderno, il potere e la conoscenza comportino grandi rischi, anche la conoscenza primitiva era pericolosa, perché nell'antichità una parte della nostra conoscenza del mondo e di noi stessi non proveniva da noi, ma dall'Altro. Dagli dèi, dagli spiriti, da voci impetuose che parlavano nel silenzio.

Al cuore dei misteri dei Veda, rivelati dal popolo dell'India, c'è l'Altare del Fuoco: una costruzione sacrificale fatta di mattoni era stata posata in precise proporzioni matematiche per assumere la forma di un enorme rapace – un'aquila, o forse un falco. Secondo Roberto Calasso era un dono fatto dalla divinità primordiale alle origini di tutto: Prajāpati, il Signore delle Creature. Quando i suoi figli, gli dèi, si lamentarono di non poter sfuggire alla Morte, gli diede istruzioni precise su come costruire un altare che permettesse loro di ascendere al cielo e ottenere l'immortalità: «Prendete trecentosessanta pietre di recinzione e diecimilaottocento mattoni, tanti quante sono le ore dell'anno», disse. «Ogni mattone ha un nome. Disponeteli in cinque strati. Aggiungetene altri fino a che siano undicimilacinquecentocinquantasei». (*Ka*, 1996) Gli dèi costruirono l'altare e sfuggirono a Mrytu, la Morte. Tuttavia, la Morte impedì agli esseri umani di fare la stessa cosa. Non ci fu concesso di diventare immortali con i nostri corpi; potevamo soltanto aspirare a opere eterne. I popoli dei Veda continuarono a erigere l'Altare del Fuoco per migliaia di anni: con il tempo, secondo Calasso, capirono che ogni mattone era un pensiero, che pensieri impilati uno sopra l'altro formavano un muro – la mente, il potere dell'attenzione – e che la mente, se sviluppata in modo adeguato, poteva volare come un uccello con le ali spiegate e conquistare i cieli.

Visti da lontano da persone che non erano a conoscenza di ciò che si stava realizzando, questi uomini e queste donne dovevano di certo sembrare dei muratori impazziti. E la stessa frenetica follia sembra possedere anche coloro che, negli ultimi decenni, hanno dedicato corpo e anima alla costruzione di un nuovo costrutto matematico, una copia senz'anima di alcuni aspetti del nostro pensiero che si è scelto di chiamare “intelligenza artificiale”, uno strumento così formidabile che, se dobbiamo credere ai più zelanti tra i suoi seguaci, ci aiuterà a raggiungere i cieli e diventare immortali.

Potere grezzo e astratto, l'IA è priva di corpo, coscienza o desiderio, e perciò, si potrebbe dire, è incapace di generare quel calore primordiale che i Veda chiamano *Tapas* – l'ardore della mente, il fervore da cui emerge l'intera esistenza – e che brucia ancora, seppur debolmente, dentro ciascuno di noi. Se ci fidiamo delle voci più ottimistiche che provengono dalla Silicon Valley, l'IA potrebbe essere il veicolo da usare per ottenere benessere illimitato, curare tutte le malattie, sanare il pianeta e spingerci verso l'immortalità, mentre i pessimisti ammoniscono che può essere la nostra rovina. È giunta la nostra ora di unirci agli dèi eterni? Oppure la nostra discendenza digitale usurperà l'Altare del Fuoco e lo adopererà per i propri fini, come noi stessi abbiamo rubato quella conoscenza, in origine destinata agli dèi? È troppo presto per dirlo. Ma di una cosa possiamo essere certi, dato che l'abbiamo appresa, più e più volte, dai racconti delle punizioni contenuti nelle nostre mitologie: non è mai sicuro esortare gli dèi, o anche soltanto avvicinarsi a loro.



A metà Ottocento, il matematico George Boole udì la voce di Dio. Mentre attraversava un campo vicino a casa sua, in Inghilterra, ebbe un'esperienza mistica e giunse a credere che avrebbe scoperto le regole soggiacenti al pensiero umano. Povero figlio di un ciabattino, Boole era un bambino prodigio che imparò da solo il calcolo e lavorò come maestro nel Doncaster finché uno dei suoi articoli non gli valse la medaglia d'oro della Royal Society, e gli garantì l'offerta di diventare il primo professore di matematica al Queen's College di Cork, in Irlanda. Con il patrocinio di un'università, e relativamente libero dalle difficoltà economiche che aveva patito così a lungo, per la prima volta poté dedicarsi quasi interamente alla sua passione, e ben presto riuscì a fare qualcosa di irripetibile: sposò matematica e logica in un sistema che avrebbe cambiato il mondo.

Prima di Boole le discipline della logica e della matematica si erano sviluppate in modo pressoché separato per oltre un migliaio d'anni. La sua nuova logica funzionava con due soli valori – vero e falso – e con essa riuscì non soltanto a praticare la matematica ma anche ad analizzare affermazioni e proposizioni filosofiche per indovinare la loro veridicità o falsità. Boole mise questo nuovo tipo di logica al servizio di qualcosa che per lui, uomo profondamente religioso, era una necessità spirituale: dimostrare che Dio è incapace di fare del male.

In una nota manoscritta che intitolò “Origine del Male”, Boole sottopose all'analisi quattro presupposti fondamentali adoperando i principi della sua logica:

1. Se Dio è onnipotente, tutte le cose devono accadere secondo la sua volontà, e viceversa;
2. Se Dio è perfettamente buono, e se tutte le cose accadono secondo la sua volontà, il male assoluto non esiste;
3. Se Dio fosse onnipotente, e se la benevolenza fosse l'unico principio della sua condotta, o non esisterebbe la sofferenza, oppure esisterebbe solamente come strumento del bene;

4. La sofferenza non esiste.

Sottopose queste affermazioni all'analisi logica sostituendole con dei simboli e combinandole in diversi modi, attraverso operazioni matematiche, finché non gli rimase un risultato che, secondo il suo sistema, era categoricamente vero: il male assoluto non esiste e la sofferenza è uno strumento del bene.

Boole era un uomo abitato dallo spirito del suo tempo, uno spirito che era molto diverso dal nostro: credeva che la mente umana fosse razionale e funzionasse secondo le stesse leggi che modellano l'intero universo; scoprendo dettagliatamente queste leggi, avremmo potuto non soltanto comprendere il mondo e rivelare i meccanismi nascosti che producono e guidano i nostri stessi pensieri, ma effettivamente scrutare anche la mente della Divinità. Dopo aver affrontato il problema del male, continuò a portare avanti le sue idee, cercando di formulare un calcolo per ridurre tutti i sillogismi, le deduzioni e le conclusioni della logica alla manipolazione di simboli matematici, e per gettare precise fondamenta alla teoria della probabilità. Tutto questo sfociò nella sua più grande opera: *An Investigation of the Laws of Thought*, un libro che definì le regole della sua nuova logica simbolica e in cui inoltre espresse, fin dal capitolo iniziale, la sua grande intenzione di catturare, con la matematica, il linguaggio di quel fantasma che mormora nei sentieri tortuosi delle nostre menti: «Il progetto del trattato che segue è di indagare le leggi fondamentali di quelle operazioni della mente attraverso cui avviene il ragionamento; per dar loro espressione nel linguaggio simbolico di un Calcolo, e su questa base stabilire la scienza della Logica e costruire il suo metodo.»

Boole era convinto che la nostra mente funzionasse sulla base di una logica che ne sta alle fondamenta, ma morì senza aver raggiunto il suo obiettivo di creare un sistema in grado di capire il pensiero; dieci anni

dopo la pubblicazione del suo capolavoro, mentre camminava nelle due miglia e mezzo che separavano la sua abitazione dall'università, fu colto dalla pioggia che lo bagnò fino a inzupparlo, quindi insegnò tutto il giorno con i vestiti fradici e gli venne un'influenza che più tardi si trasformò in una polmonite fatale. La febbre lo sprofondò nel delirio e disse a sua moglie, Mary Everest, che poteva percepire l'intero universo aprirsi di fronte a lui come un grande oceano nero, in cui non si poteva vedere né udire alcunché tranne una tromba d'argento e un coro che cantava "*Forever, O Lord, Thy word is settled in Heaven*" ("Per sempre, o Signore, la tua parola è stabile nei cieli"). Morì l'8 dicembre 1864, poco dopo che Mary (secondo una storia che potrebbe essere apocrifa) lo ebbe avvolto in lenzuola bagnate – seguendo la strana logica dell'omeopatia, secondo cui le cure devono mimare le cause – accelerando inconsapevolmente il decesso del suo amato.

Il suo lavoro non ebbe conseguenze durante la sua vita e fu ignorato per oltre ottant'anni dopo la sua morte, finché un giorno un giovane studente del MIT (Massachusetts Institute of Technology) si imbatté casualmente in *The Laws of Thought* e, immergendosi nella strana logica algebrica di Boole, creò un'applicazione pratica che da quel momento influenzò ogni aspetto delle nostre vite.

Il suo nome era Claude Shannon, matematico e ingegnere elettrico che stava lavorando sulla più sofisticata macchina pensante di quell'epoca (l'analizzatore differenziale di Vannevar Bush, un computer primordiale grande quanto una stanza), quando capì che la logica bi-valoriale di Boole era il sistema perfetto con cui progettare i circuiti elettronici. Gli interruttori elettrici adoperano valori binari (0 per spento/*off* e 1 per acceso/*on*) e possono essere controllati dalle operazioni logiche create dal matematico inglese. Si possono fare calcoli incredibilmente complessi semplicemente sfruttando una semplice dualità: vero o falso, *on/off*, 1 o 0. Questa dualità è la pietra angolare dell'Età Informatica.

Idee apparentemente inutili e altamente astratte ripresero vigore nei circuiti digitali, trasformando completamente il nostro panorama tecnologico; la stragrande maggioranza delle tecnologie che adoperano l'elettricità si basano su queste, dagli aspirapolvere ai missili balistici intercontinentali. Ma la loro applicazione più importante è il fatto che formano le basi del modo in cui i moderni computer "pensano". Questi computer "parlano" tra loro in *booleano*, e adoperano questo linguaggio per calcolare; ogni singolo compito che svolgono si riduce a una serie di domande aventi come risposte "sì o no" e che sono processate adoperando la logica di Boole. Da questa logica dipendono tutti i software e ogni singola parola del codice dei computer. La strana logica di Boole è l'alchimia invisibile che alimenta il mondo moderno, l'argilla di cui sono fatti i mattoni del nuovo Altare del Fuoco, poiché sta alle fondamenta di quell'unità di base che ispira gli odierni sistemi di IA più avanzati: il neurone artificiale.



Nel 1943 Warren Sturgis McCulloch e Walter Pitts pubblicarono il primo modello matematico di un neurone. Era estremamente semplificato e astratto, privo di ogni complicato processo della biologia reale, ma da quella semplicità derivò un enorme potere. McCulloch, un neurofisiologo e uno dei fondatori della cibernetica, e Pitts, giovane e brillante poli-matematico che eccelleva in logica, si sono basati sul fatto che, in sostanza, il comportamento di un neurone è binario: quando è stimolato, emette un impulso elettrico oppure no. Partendo da questa premessa, il loro fondamentale articolo “*A Logical Calculus of the Ideas Immanent in Nervous Activity*” (“Un calcolo logico delle attività intrinseche all’attività nervosa”) adoperò uno scalpello di Boole per forzare il funzionamento interno del meccanismo neuronale. Secondo il loro schema, ogni neurone artificiale riceve molteplici segnali elettrici dai suoi vicini, proprio come fa quello biologico; in entrambi i casi, se questi segnali superano una certa soglia il neurone si attiva, altrimenti rimane inattivo. Dopo aver creato questo costrutto matematico, fecero fare un passo in avanti alle loro idee e dimostrarono che, dal momento che sia l’*input* sia l’*output* di un neurone seguono i principi di Boole, mettendo assieme queste unità binarie in concatenazioni e circuiti, una rete costruita adoperandole avrebbe potuto calcolare ed eseguire ogni possibile operazione di logica di Boole. Da questo modello derivò una nuova comprensione del cervello e della mente; visto dalla loro prospettiva, il cervello poteva essere considerato come un dispositivo di elaborazione, una macchina che adoperava neuroni per eseguire operazioni logiche. Dunque, l’attività mentale umana non era altro che informazione binaria processata da neuroni secondo regole matematiche. Prima di McCulloch e Pitts nemmeno Turing aveva pensato di adoperare la nozione di computazione per creare una teoria della mente. Ma se il loro modello dimostrò che i neuroni artificiali sono in grado di simulare complessi processi cognitivi, si rivelò essere troppo limitato per catturare tutta l’intricata complessità delle reali menti biologiche. Ciononostante rappresentava un’intuizione enorme, perché offriva la prima teoria com-

putazionale della mente della modernità e dava una risposta a una delle più grandi domande della neuroscienza – vale a dire, come una mente diventa intelligente. Il lavoro di McCulloch e Pitts sulle reti neurali diede il via all’approccio computazionale alle neuroscienze che portò John von Neumann a creare l’architettura logica dei moderni computer. Aprì una nuova prospettiva su come potrebbero funzionare i nostri cervelli e parve mostrare come i neuroni processano e trasmettono l’informazione. A causa delle sue conseguenze su ampia scala, la loro dimostrazione che le reti neurali possono fare logica può essere considerata una delle idee più importanti nella storia del pensiero umano, ma col tempo, i loro neuroni altamente idealizzati furono scartati o ignorati dagli scienziati che lavoravano per cercare di capire il cervello, e sostituiti da schemi differenti. McCulloch passò anni a cercare di sviluppare un modello pienamente meccanicistico della mente e continuò ad andare alla ricerca della logica del sistema nervoso fino alla sua morte avvenuta nel 1969, mentre Pitts – che aveva dedicato la sua vita alla convinzione secondo la quale il misterioso lavoro compiuto dalla mente umana, le nostre molte prodezze e deficienze psicologiche, avevano origine unicamente nella meccanica dei neuroni che emettono impulsi elettrici nel cervello – cadde nella depressione e nell’alcolismo, soffrì di delirium tremens, convulsioni ed episodi di perdita di coscienza, e morì a causa di emorragie delle varici esofagee, una condizione associata alla cirrosi, da solo nella sua pensione a Cambridge, Massachusetts, dopo aver dato alle fiamme il suo lavoro su modellini tridimensionali di reti neurali. Tuttavia i neuroni artificiali sopravvissero loro e diedero la scintilla a un approccio all’apprendimento informatico che, circa quattro decenni dopo quell’articolo fondamentale, fu sostenuto con tenacia e fervore, contro le credenze comuni dell’epoca, nientemeno che da un pro-pro-pronipote di George Boole, Geoffrey Hinton.

Hinton è generalmente considerato il padrino dell'Intelligenza Artificiale. Forse è una delle singole persone che nei decenni passati ha avuto la maggiore influenza in questo campo. Negli anni '80 promosse un approccio basato sulle reti neurali profonde, astrazioni matematiche del cervello in cui i neuroni sono rappresentati con un codice; semplicemente alterando la forza delle connessioni tra i neuroni – cambiando il numero adoperato per rappresentarle – la rete poteva imparare da sé. Prima di lui il paradigma dominante era parecchio diverso: la maggior parte dei ricercatori credevano che, affinché le macchine potessero pensare, avrebbero dovuto mimare il modo in cui gli esseri umani ragionano, manipolando dei simboli (parole o numeri, ad esempio) che seguissero regole logiche, ovvero ciò che credeva lo stesso Boole. Ma il suo discendente non era d'accordo: «I corvi possono risolvere enigmi», ha detto l'anno scorso in un'intervista rilasciata alla *MIT Technology Review*, «e non hanno linguaggio... Lo fanno cambiando la forza delle connessioni tra i neuroni nel loro cervello. Dunque dev'essere possibile imparare cose complicate cambiando la forza delle connessioni in una rete neurale artificiale.»

Per molto tempo le reti neurali di Hinton non sono potute nascere. Semplicemente non avevano a disposizione una potenza di calcolo o una mole di dati su cui allenarsi tali da poter dimostrare intelligenza. Ma le cose sono cambiate in modo repentino. All'inizio degli anni 2010 ha visto sbocciare le sue idee in modi che non avrebbe potuto immaginare, e le reti neurali sono diventate il principale obiettivo della ricerca internazionale. «Abbiamo smesso di essere la frangia marginale dei matti», ha detto. «Ora siamo il nucleo centrale dei matti». Negli anni seguenti, sono stati sviluppati sistemi assolutamente sbalorditivi: AlphaGo ha demolito il campione del mondo di Go, Lee Sedol; AlphaFold ha previsto la forma di praticamente ogni struttura proteica; e programmi come DALL-E 2 ci offrono immagini fotorealistiche generate a partire dal puro e semplice rumore. Quindi è arrivata ChatGPT, un'IA capace di

fare talmente tante cose che Hinton credeva fossero lontane decenni, da far sorgere in lui il timore di Dio.

Nella primavera del 2023 Hinton ha abbandonato il suo impiego di vicepresidente di Google per mettere in guardia il mondo dai pericoli rappresentati dal frutto del suo ingegno.

Come ha spiegato a una conferenza annuale del MIT, l'IA si sta sviluppando troppo velocemente: «Non è da escludere che l'umanità sia semplicemente una fase di passaggio nell'evoluzione dell'intelligenza. Non è possibile far evolvere direttamente l'intelligenza digitale. Richiederebbe troppa energia e una meticolosa fabbricazione. C'è bisogno dell'intelligenza biologica affinché possa evolvere in modo da creare intelligenza digitale, e allora l'intelligenza digitale può assorbire in modo piuttosto lento tutto quel che le persone hanno mai scritto, che è quel che sta facendo ChatGPT, ma a quel punto può fare l'esperienza di avere accesso diretto al mondo e correre molto più veloce. Può ancora tenerci un po' per far funzionare le centrali elettriche. Ma dopo di che, forse non lo farà più.»

Hinton è cambiato. È passato dall'essere un evangelista di una nuova forma di ragione a un profeta di sventura. Ha detto che quel che gli ha fatto cambiare idea è stato l'aver capito che in realtà non abbiamo replicato la nostra intelligenza, ma ne abbiamo creata una superiore.

O forse è stato qualcos'altro? Qualche zona del suo inconscio gli ha sussurrato che era lui, e non il suo bis-bis-bisavolo, a essere stato scelto da Dio come scopritore dei meccanismi del pensiero? Hinton non crede in Dio, e probabilmente rinnegherà l'affermazione del suo antenato secondo cui la sofferenza è uno strumento della volontà del Signore, dato che è stato costretto a consumare tutti i suoi pasti sulle ginocchia, appoggiandosi a un cuscino come un monaco che prega all'altare, a causa di un

infortunio alla schiena che gli ha provocato dolori lancinanti. Per oltre diciassette anni non si è potuto sedere, e solamente nel 2022 è riuscito a farlo abbastanza a lungo da poter mangiare.

Hinton è categorico, i pericoli delle macchine pensanti sono reali. E non solamente gli effetti a breve termine come la perdita di posti di lavoro, la disinformazione o le armi letali autonome, ma un rischio esistenziale che alcuni considerano pura fantasia: che il nostro posto nel mondo possa essere soppiantato dall'IA. Parte di questo timore risiede nel fatto che lui crede che l'IA possa effettivamente raggiungere una sorta di immortalità, come fecero le divinità vediche. «La buona notizia», ha detto, «è che siamo riusciti a capire come costruire cose immortali. Quando un pezzo di hardware muore, queste non muoiono. Se avete salvato i *weights** su qualche supporto e riuscite a trovare un altro pezzo di hardware che possa eseguire gli stessi comandi, allora potete riportarlo in vita. È così che otterremo l'immortalità, ma non per noi.»

Hinton pare aver paura di quel che si potrebbe scorgere quando le braci dell'Altare del Fuoco si spegneranno alla fine del sacrificio e la pungente freddezza degli esseri che abbiamo evocato comincerà a penetrare nelle nostre ossa. Siamo davvero destinati all'obsolescenza? L'umanità peri-

* *Weight*, letteralmente “peso”; il concetto, che è l'analogo della sinapsi nelle reti neuronali biologiche, ha un ruolo fondamentale nella capacità delle reti neurali artificiali di apprendere e fare previsioni. Si tratta dei parametri che vengono regolati durante il processo di formazione per ridurre al minimo la differenza tra il risultato effettivo e quello target, valori numerici associati alle connessioni tra neuroni (o nodi) attraverso diversi strati della rete. Ogni connessione da un neurone all'altro ha un *weight* associato che indica la forza e la direzione (positiva o negativa) dell'influenza che un neurone ha su un altro. Quando un segnale in ingresso passa attraverso la rete, viene moltiplicato per questi *weight*, che determinano cumulativamente l'output finale della rete.

rà, non a causa di come trattiamo tutto ciò che ci circonda, non a causa di un'imprevedibile enorme roccia scagliata contro di noi dalla gravità, ma come conseguenza della nostra volontà irrazionale di conoscere tutto quel che può essere conosciuto? L'ipotetica apocalisse dell'IA è diversa dall'orrore della guerra nucleare che assume le sembianze di una nuvola a forma di fungo, ed è dissimile dalle devastazioni provocate da incendi, siccità e inondazioni che stanno diventando sempre più ordinarie, perché nasce da cose che, fin dall'inizio della civiltà, abbiamo sempre considerato come positive e fondamentali nel renderci umani: la ragione, l'intelligenza, la logica e la capacità di risolvere problemi, enigmi e quei mali che corrompono con sofferenze quotidiane l'esistenza anche della persona più fortunata. Ma percorrendo il nostro cammino verso l'apoteosi, osando seguire le orme delle divinità vediche che riuscirono a sfuggire alla Morte, possiamo gettare luce su cose che dovrebbero rimanere al buio. Perché anche se l'intelligenza artificiale non sarà mai all'altezza degli incubi enormi e terrificanti che fanno presagire un mondo disumano in cui gli algoritmi ronzano senza di noi, dovremo comunque fare i conti con le miriadi di effetti che questa tecnologia avrà sulla società, la cultura e l'economia degli umani.

Al tempo stesso, l'enorme spettro dell'IA superintelligente incombe su di noi. E per quanto sia poco probabile o forse perfino impossibile (nient'altro che una favola, alcuni dicono, un racconto dell'orrore che serve ad attrarre più soldi e investimenti presentando una serie di potenti sistemi non come un ulteriore passo nel nostro sviluppo tecnologico ma come un dio della morte che pone fine al mondo), non può essere scacciata tanto facilmente, perché arriva nel profondo e tocca le fibre dell'apparato con cui creiamo i miti, quella parte del nostro essere che è atavica e impaurita, perché ci ricorda un tempo in cui tremavamo nelle caverne e ci rannicchiavamo insieme, mentre fuori nell'oscurità, con occhi che possono vedere nella notte, le molte bestie selvatiche e i mostri del passato si aggiravano annusando per trovare tracce del nostro odore.

Mentre ogni nuovo modello di IA diventa sempre più forte, come le voci di un monito che provengono da un coro, e anche i più ottimisti tra di noi cominciano a temere questa nuova tecnologia, è sempre più difficile pensare senza provare panico oppure ragionare con logica. Fortunatamente, possediamo molti altri talenti che non rispondono alla ragione. E possiamo sempre alzarci, fare un passo indietro e allontanarci dal vuoto verso cui ci siamo lanciati con così tanta fretta, prestando ascolto alle strane voci che provengono dalla nostra immaginazione, quel territorio selvaggio che rimarrà sempre un rifugio necessario e un contraltare alla razionalità.

Di fronte, come siamo, alla speculazione selvaggia, dovendo confrontarci con pericoli che nessuno, per quanto intelligente e ben informato, è davvero in grado di gestire o capire, e stuzzicati dalle promesse di un potenziale illimitato, forse dovremo sondare il futuro non solamente adoperando la scienza, la politica e la ragione, ma con quell'occhio diabolico che di solito adoperiamo per vedere nelle tenebre: la narrativa. Perché nei mondi che certi autori hanno immaginato nel passato, possiamo trovare chiavi per porte che dobbiamo ancora incontrare. Mentre procediamo brancolando storditi, maltrattati e disorientati dalle capacità dell'IA, non sarebbe una cattiva idea pensare al pianeta deserto dove i



protagonisti del romanzo di Herbert, *Dune*, cercano di scrutare le onde di sabbia del futuro, in preda all'incantesimo inebriante di una droga chiamata spezia, per trovare il Sentiero Dorato, una via che permetta agli esseri umani di liberarsi dalla tirannia ed evitare l'estinzione o la stagnazione, ed essere più diversi, resilienti e liberi, evolvendoci oltre il ragionamento puramente logico e sviluppando le nostre menti e facoltà al punto che i nostri pensieri e le nostre azioni siano imprevedibili e non limitate da statistiche. I libri di Herbert, con il loro strano miscuglio di passato e presente, ci ricordano che esistono molti modi con cui poter andare avanti preservando la nostra umanità. L'IA è già arrivata, ma quel che decidiamo di farci e quali limiti siamo d'accordo a porre al suo sviluppo sono decisioni ancora da prendere. Per quanto si investano miliardi di dollari nelle compagnie di IA che promettono di eliminare il lavoro, risolvere la crisi climatica, curare il cancro e far piovere miracoli a differenza di qualunque cosa abbiamo mai vista prima, non potremo mai abbandonarci del tutto a queste creature matematiche, questi esseri senz'anima né compassione, perché non sono né vivi né coscienti – almeno non ancora, e di certo non come noi – perché non condividono con noi la natura contraddittoria delle nostre menti.

Nei prossimi anni, mentre le persone armate di IA continueranno a rendere il mondo più veloce, strano e caotico, dovremo fare tutto il possibile per evitare che questi sistemi possano dare ancor più potere a chi è in grado di costruirli. Ma dobbiamo anche prendere in considerazione un monito di Herbert, il comandamento centrale che scelse di custodire nel fondamentale testo religioso dell'umanità del futuro, una regola per impedirci di diventare asserviti ai prodotti della nostra ragione e di inchinarci di fronte al Dio della Logica e alla sua numerosa, spaventosa progenie: *Non costruirai una macchina a somiglianza della mente di un uomo.*

Pubblicato su *Harper's Magazine*, luglio 2024

ALCUNI ACULEI SPARSI DA ISTRIXISTRIX



- IX69** – TomJo, ECOLOGISMO E TRANSUMANISMO. Dei rapporti contro natura, gen17
- IX87** – Finrrage, LA LOTTA SULLE NUOVE TECNOLOGIE RIPRODUTTIVE, giu19
- IX89** – Pierrette Rigaux, LA TEORIA DEL PEZZO-DI-CARNE e SABOTARE LA MACCHINA DUALISTA, set19
- IX100** – Agustín García Calvo, CONTRO IL PROGRESSO - CONTRO IL FUTURO, nov20
- IX104** – Miguel Amorós, I TRANELLI DELL'IDENTITÀ, giu21
- IX106** – Franco Cantù / Anonimo francoprovenzale, ESTREMA (SUSS) UNZIONE / È L'ORA DELLE MEDICINE, set21
- IX107** – Franco Cantù, CONTROINDAGINE: APPUNTI SPARSI E INTEGRAZIONI SU IMPERIALISMO DIAGNOSTICO E DITTATURA TECNOSANITARIA, set21
- IX108** – Aa Vv, GREEN PASS ALL'EPOCA DELLA BIOTECNO-CRAZIA: QUALCHE VOCE CRITICA, ott21
- IX109** – Nautilus / Leonardo Lippolis, SMART CITY – L'IRRESISTIBILE ASCESA DELLA METROPOLI ECOTECNOLOGICA, ott21
- IX110** – Leonardo Lippolis, IL MONDO FUORI DAI CARDINI: NULLA SARÀ PIÙ COME PRIMA, nov21
- IX112** – Nicolas Le Dévédec, RITORNO AL FUTURO TRANSUMANISTA, nov21
- IX115** – Bernard Charbonneau / Jacques Ellul, IL PROGRESSO CONTRO L'UOMO, feb22
- IX116** – Alexandre Grothendieck, SCIENTISMO: LA NUOVA CHIESA UNIVERSALE, apr22
- X118** – Alexandre Grothendieck – CONTINUEREMO LA RICERCA SCIENTIFICA?, dic22
- IX119** – Miguel Amorós – GEOGRAFIA DELLA LOTTA SOCIALE: CAPITALISMO E CRISI ENERGETICA, gen23
- IX121** – Miguel Amorós – RITONO ALLA ANORMALITÀ AI TEMPI DEL CAPITALISMO PANDEMICO, mar23
- IX122** – Darren Allen – LA SINISTRA CONFINATA, apr23

Benjamin Labatut

THE GODS OF LOGIC
Before and after artificial intelligence



ISTRIXISTRIX@AUTOPRODUZIONI.NET

ISTRIXISTRIX.NOBLOGS.ORG

NESSUNA PROPRIETÀ

IX125

F.I.P. VIA S. OTTAVIO 20 - TORINO

LUGLIODUEMILAVENTIQUATTRO

