



Euro 26,00

ISBN 979-12-80068-53-8



ISSN 2724-0576

nuova lettera matematica

8

scienza
express

nuova

lette ra mate matica



8

**INTELLIGENZA
ARTIFICIALE** dossier
a cura di Enrico Rogora

UNA STORIA VISSUTA: IL CASO ELEA 9003 di Renato Betti / **IL
CONCETTO TOPOLOGICO E GEOMETRICO DI FORMA** di Patrizio
Frosini / **RICERCA OPERATIVA** di Gabriella Colajanni, Alessandro
Gobbi, Marinella Picchi, Alice Raffaele, Eugenia Taranto / **VITO
VOLTERRA, MATEMATICA DEL XX SECOLO (E NON SOLO) PER LE
SCUOLE SUPERIORI** di Annalisa Malusa / **DIMOSTRAZIONI: IERI,
OGGI, E DOMANI?** di Gilberto Bini

aprile 2025

l'intelligenza artificiale nell'ultimo secolo

Introduzione: il neurone e la sinapsi

In questo articolo proveremo a ripercorrere alcuni momenti salienti nello sviluppo dell'intelligenza artificiale (IA), evidenziando come molteplici branche della matematica si siano alternate, e si alternino tutt'oggi, in maniera sinergica per contribuire all'impalcatura di un telaio logico nel quale razionalizzare questa disciplina. Inoltre, ci proponiamo di analizzare i punti di forza e le criticità del primo tentativo di sviluppo dell'IA, realizzato nella prima metà del secolo scorso, e di quello successivo, avviato negli anni Ottanta e ancora in corso.

Sancire un momento di nascita della computazione automatica è arduo se non impossibile: le origini sono complesse e affondano le radici in ambiti diversi, basti pensare che la macchina di Babbage, un primordiale calcolatore, traeva ispirazione dalle automazioni sviluppatesi nell'industria tessile alla fine del Settecento, come il telaio di Jacquard.

Un momento che potrebbe segnare un inizio ragionevole della storia



Elena Agliari

Laureata in Fisica all'Università di Parma, è professoressa associata in Fisica Matematica presso il Dipartimento di Matematica di Sapienza Università di Roma, dove insegna Modelli di reti neurali e Matematica per il machine learning. Si è occupata di processi di diffusione su grafi, modelli matematici applicati alle reti immunitarie e di meccanica statistica dei sistemi disordinati.

elena.agliari@uniroma1.it



Adriano Barra

Laureato in Fisica (Teorica) presso Sapienza Università di Roma e dottorato in Matematica (Applicata) presso King's College University of London, è professore associato in Fisica matematica presso il Dipartimento di Scienze di base applicate all'ingegneria della Sapienza Università di Roma, dove tiene un corso di reti neurali e machine learning per il dottorato. La sua ricerca verte prevalentemente su meccanica statistica (di reti neurali e sistemi disordinati) e modellizzazione matematica di fenomeni biologici.

adriano.barra@uniroma1.it

moderna di questa disciplina è costituito dal chiudersi della *Belle Époque*, quando si raggiunse una soddisfacente comprensione del campo elettromagnetico: questo permise un notevole progresso nella fisiologia (in particolare la branca che oggi prende il nome di neuroscienze) e nella sua controparte artificiale (composta da diversi settori affini quali robotica e automazione, cibernetica e teoria dell'informazione), da cui nasceranno le prime teorie per le reti neurali. La genesi di tale progresso è semplice a capirsi: mentre molte reti del nostro corpo sono reti chimiche, i neuroni dialogano tra loro inviandosi impulsi elettrici. In questo contesto, un contributo fondamentale è dovuto a Camillo Golgi e Santiago Ramon y Cajal (entrambi premi Nobel per la Medicina nel 1906), che, a cavallo tra il XIX e il XX secolo, svilupparono nuove tecniche sperimentali grazie alle quali divenne finalmente possibile mettere a fuoco la struttura istologica del sistema nervoso e iniziare a studiarne i componenti, ossia i neuroni¹. In seguito, una vasta schiera di scienziati, tra cui Ian Lloyd Hodgkin e Andrew Fielding Huxley (entrambi premi Nobel per la Medicina nel 1963), si dedicò a queste ricerche, riuscendo a spiegare i meccanismi con cui i neuroni comunicano tra loro, ovvero come generano e trasmettono impulsi elettrici lungo le connessioni che li collegano (Tuckwell 1988). Semplificando, ogni neurone ha un corpo (soma) isolato dall'esterno tramite una membrana di fosfolipidi e vive mantenendo una differenza di potenziale tra il suo interno e l'ambiente esterno (Figura 1). Gli stimoli provenienti dai neuroni ad esso adiacenti (o eventualmente dal mondo esterno se osserviamo neuroni periferici, come i neuroni retinici) concorrono a incrementare o decrementare (a seconda che la connessione sia eccitatoria o inibitoria, come spiegheremo nel seguito) tale differenza di potenziale e, se questa raggiunge una soglia critica, l'intera stabilità cellulare viene meno e il neurone "spara" un segnale elettrico lungo il suo assone per poi ripristinare la sua differenza di potenziale al livello basale. Il segnale elettrico emesso, detto *spike*, viaggia lungo l'assone

1 Per renderci conto di quanto recente sia la nostra (comunque acerba) comprensione delle reti neurali biologiche è interessante osservare che il termine *neurone* venne coniato da W. Waldeyer-Hartz solamente nel 1891.

dossier a cura di Enrico Rogora

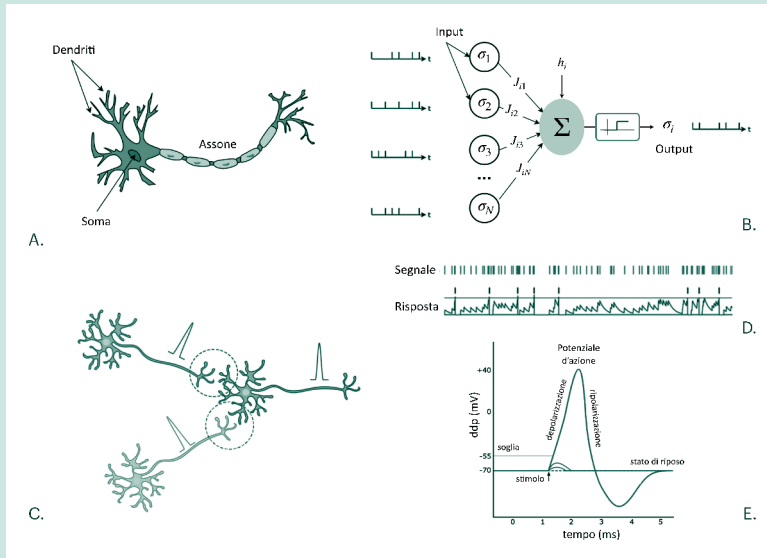


Figura 1. Rappresentazione minimale di un neurone biologico (A) e di un neurone artificiale (B). Ogni neurone può emettere un segnale elettrico che si propaga lungo il relativo assone, connesso tramite i bottoni sinaptici ai dendriti dei neuroni adiacenti (C). Un generico neurone, σ_i , è connesso ad altri N neuroni, ciascuno con una propria attività (come indicato in figura dai relativi “treni di spike”), che complessivamente forniscono al neurone i -esimo un segnale (D), tale segnale può dar luogo ad un’attivazione del neurone i -esimo, secondo un meccanismo a soglia (E).

e influisce a sua volta sulla differenza di potenziale di altri neuroni a questo collegati. Se consideriamo una rete composta da N neuroni e schematizziamo la risposta del singolo neurone come un sistema *on/off* a due stati, in maniera tale che, preso il neurone i -simo, $\sigma_i = 1$ rappresenti il suo stato attivo e $\sigma_i = 0$ il suo stato quiescente, la funzione di risposta di questa cellula può essere schematizzata in modo estremamente efficace, come proposto da Walter Pitts (matematico) e Warren McCulloch (neurologo) nel 1943 (McCulloch & Pitts 1943), i quali descrissero questi neuroni come variabili di Boole attivati da funzioni Θ di Heaviside:

$$\sigma_i(t + 1) = \theta \left(\sum_{j=1}^N J_{ij} \sigma_j(t) - h_i \right),$$

dove t indica il tempo discretizzato, h_i rappresenta l’eventuale presenza di stimoli esterni e la soglia di attivazione è stata fissata (senza perdita di generalità) al valore zero. Le giunzioni

$J_{ij} \in \mathbb{R}$ tra due neuroni – nello specifico tra l'assone terminale di un neurone j e le dendriti (cioè le parti dedite a ricevere gli *spike* afferenti) di un neurone i – sono chiamate *sinapsi*: queste, mediante il rilascio di neurotrasmettitori, possono passare al neurone adiacente un messaggio eccitatorio ($J_{ij} > 0$) o un messaggio inibitorio ($J_{ij} < 0$) mediante opportuni neurotrasmettitori (e.g., serotonina e dopamina o GABA e glicina, rispettivamente). Quindi, in questa rete neurale, sia i neuroni (pensati come i nodi della rete) che le sinapsi (pensate come gli archi tra tali nodi) rivestono un ruolo cruciale e le relative variabili sono soggette a dinamiche specifiche, con scale di tempo caratteristiche, da cui emerge la capacità della rete di elaborare informazioni. In particolare, la dinamica delle sinapsi è lenta e riguarda l'apprendimento (la memorizzazione di un'immagine, per fare un esempio semplice), mentre quella dei neuroni è veloce e concerne il riconoscimento (dell'immagine precedentemente appresa). Infine, notiamo che, poiché la funzione di risposta del singolo neurone descritta dall'equazione è alquanto elementare, le capacità del cervello (come per esempio memorizzare o riconoscere immagini) devono emergere come *proprietà collettive* della rete.

Il primo programma per l'AI

Queste prime nozioni di neuroscienze, con le relative modellizzazioni matematiche, portarono alla creazione dei primi prototipi di neurone artificiale, tra cui il celebre *perceptrone* introdotto nel 1957 da Frank Rosenblatt (Rosenblatt 1958). Questa primitiva macchina per l'apprendimento imitava, mediante componenti elettronici, le conoscenze neuroscientifiche disponibili all'epoca, ed è interessante notare come, già allora, l'interesse per questa IA (sebbene ancora rudimentale) fosse altissimo al punto che della ricerca di Rosenblatt ne parlò, con toni sensazionalistici, anche il New York Times. Proprio l'anno precedente, il 1956, è ricordato come l'inizio ufficiale dello studio dell'IA, a seguito del convegno dedicato a questo tema tenutosi presso il Dartmouth College. Numerosi scienziati del calibro di Marvin Minsky e Claude Shannon parteciparono con l'obiettivo di progettare macchine capaci di simulare ogni aspetto dell'apprendimento e dell'intelligenza umana.

Per un lettore matematico è forse interessante esaminare anche l'evoluzione dell'impiego di diversi strumenti matematici nello sviluppo di una teoria per l'IA. In particolare, nella prima metà del secolo scorso, periodo in cui venne avviato il primo programma di ricerca, la visione dominante era fortemente influenzata da una logica meccanicistica. Il programma di David Hilbert dominava la scena, mentre i *Principia Mathematica* di Alfred Whitehead e Bertrand Russell, pubblicati in tre volumi tra il 1910 e il 1913, lasciavano un'impronta profondissima. Tra il 1930 e il 1931, Kurt Gödel formulò i suoi teoremi di completezza e incompletezza, anticipando di poco i contributi di un altro gigante del primo programma, John von Neumann. Nello stesso periodo, Alan Turing pubblicava il suo celebre lavoro *On Computable Numbers, with an Application to the Entscheidungsproblem*, in cui descriveva per la prima volta la futura "macchina di Turing". Contemporaneamente, Claude Shannon e il suo relatore di dottorato Norbert Wiener gettavano le basi della prima teoria cibernetica.

Non sorprende, dunque, che in quel periodo l'intelligenza fosse concepita come una proprietà fondata su rigorosi ragionamenti deterministici, imitabile tramite operazioni logiche. In questo contesto, il calcolo proposizionale, grazie alla sua implementazione circuitale che permetteva di eseguire concretamente le funzioni booleane, appariva perfettamente adeguato. L'idea centrale era quella di scrivere programmi deterministici, capaci di formalizzare progressivamente compiti sempre più complessi, eseguibili da macchine progettate come reti di porte logiche. Questo approccio guadagnò terreno fino a incontrare una crisi significativa, segnata dalla critica al perceptrone avanzata dai matematici Marvin Minsky e Seymour Papert nel loro celebre libro *Perceptrons* (Minsky & Papert 1969). In quest'opera, gli autori mettevano in luce i limiti profondi della macchina di Rosenblatt, evidenziando, tra le altre cose, come il perceptrone non fosse neppure in grado di funzionare come una porta logica XOR. In un mondo guidato dal meccanismo del calcolo proposizionale, questo "difetto" mostrava che difficilmente tali sistemi avrebbero potuto dare origine alle meraviglie dell'IA di cui siamo testimoni oggi. Un'altra

palese dimostrazione dei limiti di questo approccio si ebbe sul finire del secolo scorso, quando per la prima volta un computer – il *Deep Blue* di IBM – sconfisse il campione del mondo di scacchi in carica, Garry Kasparov. Questo evento provò al mondo la straordinaria complessità logica delle macchine di Turing, culmine del primo programma per l'IA. Tuttavia, la stessa classe di processori di impostazione tradizionale si rivelava enormemente inadeguata – con risultati spesso modesti – in compiti considerati banali per un cervello umano, come il riconoscimento di oggetti in una fotografia (ad esempio, una sedia in un salone). Questo tipo di attività, noto come *pattern recognition*, sarebbe invece diventato il fulcro del secondo programma per l'IA, alla base delle moderne applicazioni come i veicoli a guida autonoma.

In sintesi, i ricercatori si trovarono di fronte a una realtà inaspettata e alla consapevolezza che scrivere un programma abbastanza versatile da simulare funzioni cognitive elementari, come il riconoscimento di categorie (ad esempio, identificare l'immagine di una sedia), fosse proibitivamente difficile: imitare il cervello esclusivamente attraverso protocolli deterministici implementati su reti di porte logiche rappresentava, al massimo, solo una parte della storia. Quest'*impasse* indusse la ricerca a esplorare nuove vie, ponendo le basi per il secondo programma dell'IA.

Il secondo programma per l'IA

Gli anni Settanta del secolo scorso rappresentano un periodo di stagnazione nella comprensione delle reti neurali noto come *winter time*². D'altra parte, tale periodo permise una profonda riflessione che portò alla consapevolezza delle criticità del primo programma per l'IA e condusse al secondo programma per l'IA, che continua fino ai nostri giorni. A

² Il *winter time* inizia negli anni Settanta (oltre alla critica formale al perceptrone, il programma di ricerca della DARPA – ente della Difesa Americana – sulla comprensione del linguaggio arrancava) e prosegue negli anni Ottanta quando, in conseguenza del decennio precedente, il governo americano ridusse sensibilmente i fondi destinati a queste ricerche e, in ragione del *Lighthill report*, lo stesso avvenne in Inghilterra, per poi rimanere fluttuante fino al decennio del 2010, noto come *spring time* (Bommasani 2023). Si veda anche (Chaitin 2009) per alcuni aspetti algoritmici.

nostro avviso, i punti di svolta maturati in quel decennio possono essere identificati in due aspetti fondamentali.

Dal punto di vista del *software* (o, in termini moderni, del *machine learning*) si accantonò progressivamente l'idea di scrivere codici deterministici (dove era la bravura del programmatore ad astrarre l'archetipo di sedia e definirlo operativamente per la macchina), per adottare invece un approccio basato sull'uso di numerosi esempi della categoria da astrarre (in questo caso, la sedia) da fornire alla macchina, che le permettessero così di apprendere autonomamente. In questo modo, il calcolo si spostò da una visione deterministica a una stocastica, fondata sulla statistica e sull'inferenza.

Dal punto di vista dell'architettura, cioè dell'*hardware* (o, in termini moderni, delle *neural networks*), una progressiva e silenziosa rivoluzione spostò l'attenzione dai soggetti alle loro interazioni: si comprese infatti che, mentre il perceptrone di Rosenblatt non poteva agire come una porta logica universale, un insieme di perceptron poteva farlo. Di conseguenza, il focus non era più sul singolo neurone, ma sulle interazioni che scaturivano da un'assemblea di neuroni e che consentivano di prendere decisioni collettive; tale cambiamento portò alla necessità di comprendere i fenomeni emergenti all'interno di reti di elementi in interazione. Notiamo che questo cambio di prospettiva, dal soggetto alle sue interazioni, è tutt'altro che banale. Ad esempio, il copernicano abbandono del modello antropocentrismo rappresenta, a ragione, una svolta epocale, ma sposta il focus da un soggetto (la Terra) a un altro (il Sole) senza porre in rilievo le interazioni gravitazionali tra le masse. Inoltre, con l'avvento della moderna teoria dei grafi – avviata dai matematici Paul Erdős e Alfréd Rényi verso la fine degli anni Cinquanta – l'intera scienza moderna si è arricchita di una descrizione in termini di reti. Si pensi, ad esempio, alle reti metaboliche, immunitarie e geniche in biologia, o alle reti sociali, economiche e ai *food web* in sociologia ed ecologia, solo per citare alcuni esempi oggi ormai consolidati (Caldarelli 2007).

L'IA non è rimasta immune a questo cambio di paradigma, e i risultati derivanti da un approccio basato sulla costruzione di reti capaci di apprendere statisticamente dall'esperienza (ossia

dai dati) non tardarono a manifestarsi: le prime pietre miliari furono poste da John Hopfield e Geoffrey Hinton (entrambi premi Nobel per la Fisica nel 2024 per i loro contributi pionieristici alle reti neurali) e risalgono agli anni Ottanta del secolo scorso. Da questo momento in poi, decennio dopo decennio, le metodologie tradizionali si consolidarono in alcune branche dell'ingegneria, dell'automazione e dell'informatica (dove trovarono una felice applicazione e continuano tutt'oggi a occuparsi degli aspetti più meccanicistici della teoria), mentre, oltre alla nuova teoria dei grafi, anche l'inferenza statistica (in particolare nella sua declinazione ad alta dimensionalità³) e la meccanica statistica (in particolare quella dei sistemi complessi (Mézard *et al.* 1986, Coolen *et al.* 2005) vennero progressivamente cooptate nel campo dell'IA.

Nello specifico, nel 1982 John Hopfield (Hopfield 1982) implementò i meccanismi di apprendimento formulati empiricamente da Donald Hebb (Hebb 1949) (che a sua volta ereditò e sintetizzò gli studi sui riflessi condizionati condotti da Ivan Pavlov, il *Galileo* della fisiologia, alcune decadi prima (Agliari *et al.* 2023)), in una rete che è spontaneamente in grado di fare *pattern recognition* (cioè, di riconoscere una sedia in un'immagine) e nella quale la memoria non risiede in un luogo altro rispetto a quello dove avviene la computazione (come nei computer, dove il disco fisso è fisicamente separato da CPU o GPU), ma è distribuita tra le sinapsi e funziona in maniera associativa, proprio come nel cervello. L'idea seminale di Hebb, racchiusa nella celebre frase *neurons that fire together wire together*, è strabiliante nella sua semplicità: di fronte alla persistenza di un nuovo stimolo (ad esempio, durante l'apprendimento di una nuova immagine), la regola di apprendimento hebbiana prevede che i neuroni che si scambiano messaggi elettrici in modo continuativo rafforzino le loro connessioni reciproche come conseguenza. Se si considera la corrente come un flusso (di elettroni), questa regola diventa un'elegante prescrizione di ottimizzazione idraulica: dove passa un grande flusso, il condotto deve essere ampliato;

3 Questo termine si riferisce a un ambito dell'inferenza che, trascurato nel secolo scorso, si occupa di modelli il cui numero di parametri cresce proporzionalmente al numero di dati.

al contrario, poiché le risorse sono limitate, i condotti poco utilizzati devono essere ristretti⁴.

Diverse sono invece le reti neurali di Hinton, su cui spenderemo una manciata di parole prima di avviarci alle conclusioni: una volta impostata una prospettiva statistica, non tutti i ricercatori sentivano l'esigenza di aderire alla controparte biologica e, nella nascente *Computer Science*, alcuni di loro si concentrarono su algoritmi di apprendimento e annessi architetture neurali che potessero garantire prestazioni ottimali, a prescindere dalla loro plausibilità biologica. Se Hopfield è stato il capostipite del filone di ricerca *bio-inspired*, probabilmente Hinton ha avuto lo stesso ruolo nel ramo della ricerca *data-driven*, per cui non possiamo non menzionare la *backpropagation* per le reti *feed-forward* (LeCun *et al.* 1988) e la *contrastive divergence* per le macchine di Boltzmann (Hinton *et al.* 1985) come due tra i più popolari algoritmi per l'apprendimento automatico. In particolare, il primo – utilizzato estensivamente nell'apprendimento supervisionato di reti profonde (LeCun *et al.* 2015) – serve a valutare come modificare i pesi delle connessioni sinaptiche, mediante retro-propagazione dell'errore, nel calcolo del gradiente di una funzione costo in uno spazio ad altissima dimensionalità, una tecnica basata sull'analisi matematica e il calcolo numerico più che su meccanismi d'ispirazione biologica⁵.

Conclusioni

In questo scritto abbiamo cercato di evidenziare come questa IA di cui tanto si parla oggi sia una disciplina con una lunga storia e, tralasciando per ovvi motivi di spazio

4 Per il suo valore storico, e considerando che questo testo divulgativo è rivolto a un lettore italiano, è interessante notare che una prima trattazione sistematica del modello di Hopfield tramite la meccanica statistica di Parisi è dovuta a Daniel Amit e collaboratori (Amit 1989, Amit *et al.* 1985), e la dimostrazione matematica rigorosa della correttezza della teoria di Parisi si è conclusa nel 2003, grazie a un *tour de force* tra Francesco Guerra (Guerra 2003) e Michael Talagrand (Talagrand 2006), quest'ultimo insignito del premio Abel nel 2023 anche per la sua dimostrazione della formula di Parisi. Rimandiamo a (Agliari & Barra 2020) per una nostra trattazione divulgativa di questi temi.

5 Solo negli ultimi anni, si è dimostrato come, in realtà, la macchina di Hinton (nota come *Boltzmann machine*) abbia legami strettissimi con la rete di Hopfield e in alcuni semplici casi analizzabili analiticamente, queste siano addirittura equivalenti (Barra *et al.* 2012) e perfino la *backpropagation* possa trovare una controparte biologica (Hinton *et al.* 2020).

il suo passato remoto, abbiamo posto enfasi sui due principali programmi di ricerca in IA che si sono succeduti negli ultimi cento anni. Lo scopo non è raccontare qualche sbalorditivo successo, ma piuttosto condividere con il lettore l'idea che questi risultati, oltre a essere l'occasionale trionfo di una rete (o di un algoritmo), verosimilmente dimenticato in una manciata di mesi poiché oscurato dal suo successore, sono il frutto del lavoro di molte generazioni di scienziati, che si sono succedute sinergicamente attingendo alle discipline più disparate (di cui abbiamo menzionato alcuni rami della matematica dato il bacino di lettura della presente rivista, ma non mancano nel paniere contributi pivotali da psicologia cognitiva, fisica teorica, linguistica, ecc.).

In quest'ottica, abbiamo sottolineato come il primo programma per l'IA non costituisca un mero fallimento, sconnesso dal secondo programma, ma piuttosto come suo necessario precursore. Infatti, la creazione delle macchine di Turing, cioè dei nostri computer, era la *conditio sine qua non* per iniziare a “digitalizzare la realtà”, fenomeno al quale tutti contribuiamo quando condividiamo sui *social media* informazione “etichettata” (e.g., la foto di una sedia con un commento come “oggi ho comprato questa bella sedia”): tali informazioni digitali costituiscono esempi già classificati, fondamentali per addestrare reti e algoritmi (in questi esempi noi fungiamo da *teacher*, scrivendo cosa la rete deve aspettarsi di trovare nell'immagine). È chiaro dunque che, nonostante le prime teorie sulle reti neurali siano state sviluppate (da Hopfield, Hinton e molti altri) negli anni Ottanta, finché la rappresentazione digitale della realtà da apprendere e da esporre a questi cervelli artificiali non era abbastanza estesa, era piuttosto illusorio sperare di osservare anche una minima complessità cognitiva – d'altra parte, anche una rete neurale biologica necessita di esperienze e stimolazioni per sviluppare capacità intellettive avanzate.

Con il secondo programma e il passaggio ad architetture sofisticate, allenate tramite tecniche stocastiche ed enormi quantità di dati, si sono raggiunte capacità sorprendenti, in alcuni casi addirittura superiori a quelle di un essere umano esperto, tanto da parlare di una vera e propria rivoluzione e, come in ogni rivoluzione, le luci si accompagnano a ombre,

che possono destare qualche preoccupazione. Tra queste, la mancanza di un controllo puntuale su cosa effettivamente accada all'interno di queste reti neurali, che appaiono così come scatole nere: questa è la conseguenza inevitabile del passaggio da protocolli deterministici ad algoritmi *data-driven* che la ricerca moderna cerca di affrontare sviluppando la cosiddetta XAI (*eXplainable Artificial Intelligence*) (Pedreschi *et al.* 2018).

Il consumo energetico è un altro aspetto critico, basti pensare che l'addestramento di una macchina moderna (e.g., un telaio da *deep learning*) su un problema arduo (e.g., un *large language model*) può arrivare a richiedere l'equivalente consumo ed emissione di CO₂ di cinque automobili nell'arco della loro intera vita (Hao 2019). Anche lo stoccaggio di dati richiede importanti risorse e molti colossi dell'hi-tech (Microsoft, Amazon, ecc.) stanno trasferendo i loro server in fondo all'oceano, al fine di sfruttare il calore latente dell'acqua, anziché la climatizzazione canonica, per il raffreddamento. La soluzione per migliorare l'efficienza di queste macchine potrebbe non essere troppo lontana, prendendo ispirazione da una macchina particolarmente efficiente nella elaborazione delle informazioni: il nostro cervello, frutto di milioni di anni di evoluzione.

In effetti, in tempi recenti, stiamo assistendo a un ravvicinamento tra le neuroscienze e l'IA. Questa volta l'interesse non riguarda i meccanismi della circuitistica elementare (come accadde con la stilizzazione del neurone nel perceptrone, ormai ben compreso), bensì fenomeni cognitivi più complessi e strutturati. Ad esempio, negli ultimi anni sono state messe a punto reti neurali in grado di "dormire" e si è compreso come, equipaggiando queste macchine con algoritmi che simulino le varie fasi del sonno, si possa ottenere un significativo ritorno pratico. Nello specifico, abbiamo dimostrato (Agliari *et al.* 2019) come una rete di Hopfield che possa "ristorarsi" all'occorrenza, offra le stesse prestazioni di una rete di Hopfield classica (cioè perennemente operativa) arrivando però a risparmiare fino al 90% del volume di dati per il suo addestramento (e quindi, in ultima istanza, di consumo energetico). Esiste addirittura un mercato per le immagini prodotte da Dee-

pDream (una rete neurale creata dall'ingegnere di Google Alexander Mordvintsev), che vengono vendute come vere e proprie opere d'arte lasciando trasparire come, forse, nella nostra società, ci si dedichi troppo all'intelligenza artificiale e poco a quella biologica.

Bibliografia

Agliari E. *et al.*, *Dreaming neural networks: forgetting spurious memories and reinforcing pure ones*. *Neur. Netw.* 112:24-40, 2019.

Agliari E. *et al.*, *From Pavlov Conditioning to Hebb Learning*, *Neural Comp.* 35(5):930, 2023.

Agliari E., Barra A., *L'elaborazione d'informazione nelle reti neurali*, Ithaca XVI, 25-39, 2020); *La meccanica statistica dei sistemi complessi*, Ithaca XVI, 209-222, 2020.

Amit D.J., *Modeling brain functions*, Cambridge University Press, 1989.

Amit D.J. *et al.*, *Storing infinite numbers of patterns in a spin-glass model of neural networks*, *Phys. Rev. Lett.* 55(14):1530, 1985.

Barra A. *et al.*, *On the equivalence among Hopfield neural networks and restricted Boltzmann machines*, *Neur. Netw.* 34:1, 2012.

Bommasani R., *AI Spring? Four Takeaways from Major Releases in Foundation Models*, *Language Processing & Machine Learning*, Stanford University, 2023.

Caldarelli G., *Scale-free networks: complex webs in nature and technology*, Oxford University Press, 2007.

Chaitin G.J., *Algorithmic Information Theory*, Cambridge University Press, 2009.

dossier a cura di Enrico Rogora

Coolen A.C.C., Kühn R., Sollich P., *Theory of neural information processing systems*, Oxford Press, 2005.

Guerra F., *Broken replica symmetry bounds in the mean field spin glass model*, Comm. Math. Phys. 233:1-12, 2003.

Hao K., *Training a single AI model can emit as much carbon as five cars in their lifetimes*, MIT Tech. Rev., 2019.

Hebb D., *The organization of behavior*, Psychology Press, 1949.

Hinton G.E. et al., *A learning algorithm for Boltzmann machines*, Cognit. Sci. 9(1):147, 1985.

Hinton G.E. et al., *Backpropagation and the brain*, Nature Rev. Neurosci. 21(6):335-346, 2020.

Hopfield J.J., *Neural networks and physical systems with emergent collective computational abilities*, Proc. Natl. Acad. Sci. 79(8):2554, 1982.

LeCun Y. et al., *A theoretical framework for back-propagation*, Proc. Connectionist Mod., 1988.

LeCun Y., Bengio Y., Hinton G., *Deep learning*, Nature 521.7553: 436-444, 2015.

McCulloch W.S., Pitts W., *A logical calculus of the ideas immanent in nervous activity*, Bull. Math. Biophys. 5:115-133, 1943.

Mézard M., Parisi G., Virasoro M., *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications*, World Scientific Press, 1986.

Minsky M., Papert S., *Perceptrons: An Introduction to Computational Geometry*, 1969.

Pedreschi D. et al., *A survey of methods for explaining black box models*, ACM Comp. Sur. 51:(5), 1-42, 2018.

Rosenblatt F., *The perceptron*, Psychological Rev. 65.6:386, 1958.

Talagrand M., *The Parisi formula*, Annals of Mathematics, 221-263, 2006.

Tuckwell H.C., *Introduction to theoretical neurobiology: linear cable theory and dendritic structure*, Cambridge University Press, 1988.