

«SEGNATURE»

Le ragioni della scienza pura nell'era dell'Intelligenza Artificiale

Alberto QUADRIO CURZIO (Presidente Emerito dell'Accademia Nazionale dei Lincei)

Le ragioni della scienza pura nell'era dell'Intelligenza Artificiale

Angelo Maria PETRONI (Sapienza Università di Roma)

La scoperta scientifica e la sua logica nell'era dell'Intelligenza Artificiale

Paolo CASINI (Linceo, Sapienza Università di Roma)

La coscienza di Turing

Alfio QUARTERONI (Linceo, Politecnico di Milano; EPFL - Lausanne)

Come riconciliare intelligenza artificiale e scienza dei principi primi

Roberto SCAZZIERI (Linceo, Alma Mater Studiorum Università di Bologna; Gonville and Caius College e Clare Hall, Cambridge)

Intelligenza artificiale, condotta umana e teoria: una prospettiva di analisi economica

Roma, 8 novembre 2023

ALBERTO QUADRIO CURZIO^(*)

LE RAGIONI DELLA SCIENZA PURA NELL'ERA DELL'INTELLIGENZA ARTIFICIALE

1. PRESENTAZIONE

Il focus di questa «Segnatura» è la distinzione tra scienza pura e scienza finalizzata. Sin dal 1963 l'OCSE fornisce definizioni e misurazioni accurate, volte alle misurazioni degli investimenti rispettivi. Misurazioni che mostrano come gli investimenti in scienza pura tendano a diminuire in termini relativi.

L'Intelligenza Artificiale introduce una nuova dimensione. Infatti, viene spesso affermato che l'avvento dell'IA renderebbe obsoleto lo stesso concetto di metodo scientifico nato con la prima Rivoluzione Scientifica, e con esso la stessa ragion d'essere della scienza pura.

L'affermarsi di questa tesi avrebbe conseguenze notevolissime. Vi sarebbero conseguenze sul piano della stessa concettualizzazione della scienza e del suo valore. Vi sarebbero altresì conseguenze sulle politiche della ricerca scientifica nelle società democratiche contemporanee, con una verosimile riduzione delle risorse destinate alla scienza pura, e con l'indebolimento della dimensione di bene pubblico universalistico della conoscenza scientifica nel suo complesso. Il tema sarà discusso dal punto di vista delle scienze filosofiche, matematiche ed economiche.

Così questa Segnatura è stata presentata istituzionalmente e ritengo che la stessa si collochi con coerenza anche nella interdisciplinarietà.

(*) Presidente Emerito dell'Accademia Nazionale dei Lincei.

2. LA STRUTTURA: FILOSOFIA, MATEMATICA ED ECONOMIA

L'origine della segnatura è infatti connessa ai miei frequenti colloqui con un filosofo della scienza (Angelo Petroni, distaccato al Centro Linceo Segre per tre anni) e alla collaborazione con un economista strutturalista (Roberto Scazzieri, Linceo). Con loro ho parlato spesso della Intelligenza Artificiale come economista strutturalista, convinto che l'interdisciplinarietà sia essenziale soprattutto per capire come un'innovazione di questa portata possa cambiare radicalmente la dinamica dei sistemi economici. Essendo altresì consapevole che la Filosofia e la Matematica sono Scienze portanti del Sapere mentre l'Economia è una componente essenziale (necessaria ma non sufficiente) dello Sviluppo. L'economia, come altre scienze, è espressa in varie teorie e correnti di pensiero e tra queste l'Economia Strutturalista è probabilmente la più consapevole che, oltre alle Scienze Sociali, la Filosofia e la Matematica sono due riferimenti ineludibili. Così come l'Economia Strutturalista, collocandosi anche in una dimensione storica dei fatti, tiene conto che nel lungo periodo si sono spesso presentati eventi del tutto imprevedibili di fronte ai quali la "razionalità degli agenti economici e la mano invisibile del mercato" non furono bastanti per capire e per decidere. Affidarsi in una situazione come questa a loro significa affidarsi al caso. Per questo oltre ai due colleghi prima citati, con i quali ho spesso discusso di questi temi, ho ritenuto di interpellare per una valutazione complessiva del problema due Soci lincei che non erano stati partecipi dei miei colloqui con Petroni e Scazzieri e che rappresentassero, in modo del tutto indipendente, la Filosofia e la Matematica. Ringrazio quindi i Lincei Paolo Casini e Alfio Quarteroni per aver accolto il mio invito. Tuttavia, visto che l'iniziativa era incardinata sui citati colloqui, ho chiesto a Petroni di fare la relazione di apertura e a Scazzieri quella di chiusura.

Data la rilevanza e la incisività delle relazioni non intendo certo farne qui una sintesi; riprendendo invece le valutazioni conclusive di ciascun relatore che mi paiono complementari e potenzialmente utili per ulteriori collaborazioni.

3. LE RAGIONI DELLA SCIENZA PURA NELL'ERA DELL'INTELLIGENZA ARTIFICIALE: PETRONI

«Le considerazioni che abbiamo esposto sono del tutto distinte da ogni posizione che sostenga la intrinseca impossibilità di una Intelligenza Artificiale che possa replicare l'intelligenza umana. La posizione più nota in questo è probabilmente quella di John Searle e, come è noto agli storici della filosofia, la questione può essere significativamente fatta risalire a Cartesio e a Leibnitz, con argomenti che ancora oggi hanno un interesse e una rilevanza

del tutto notevoli. Personalmente riteniamo che il tentativo di *dimostrare* dei limiti ontologici, o anche solo metodologici, alla capacità di una macchina – comunque concepita o effettivamente costruita – di replicare l'intelligenza umana, non abbia maggior fondamento di quanto ne avesse all'epoca (1872), e ne abbia avuto successivamente, *l'Ignorabimus di Emile Du Bois-Raymond, con l'enunciazione dei sette "misteri" destinati per sempre a sfuggire alla spiegazione scientifica*. Per chi condivida una posizione materialistica l'idea di una mente/cervello umana come realtà unica e di principio non riproducibile, che sia da parte della natura o dell'uomo stesso, è errata. La nostra tesi per cui l'attività di "Theory building", il "nocciolo duro" della scienza pura, non può essere sostituita dall'Intelligenza Artificiale *come noi la conosciamo oggi* non ambisce e non potrebbe ambire ad affermare alcuna impossibilità di principio. Si tratta di una tesi, certo non particolarmente originale, che deriva esclusivamente da considerazioni di tipo logico e metodologico, guidate da una interpretazione di alcuni momenti significativi della storia della scienza».

4. LA COSCIENZA DI TURING: CASINI

«La meta di un archivio universale sembra ormai prossima, ma gli esiti dell'onniscienza archiviata potrebbero risultare anch'essi ambigui. In migliaia di secoli di evoluzione dall'*homo erectus* al *faber* al *sapiens* milioni di generazioni si sono adattate all'ambiente e lo hanno modificato memorizzando e modificando di generazione in generazione patrimoni teorici e pratici di cognizioni.

La fatica spesa nell'apprendimento di miliardi di umani per bilioni di ore non è da escludere dal bilancio finale. La reazione al vecchio, la curiosità del nuovo hanno agito da stimolo nelle varie rivoluzioni scientifiche. Ci si può chiedere se il feed-back dalla superintelligenza ai singoli utenti umani potrà stimolare ulteriormente questo processo, oppure favorire abulia e inappetenza del sapere. Se le macchine diventassero protagoniste della ricerca e controllassero le menti umane, potrebbero considerarci come una specie rimasta indietro nell'evoluzione, lasciarci vivere con il loro beneplacito – secondo il "Gorilla problem", un luogo comune nella letteratura sulla AI – come gli umani conservano la specie protetta del Gorilla nelle foreste

Ci si può chiedere infine come convivere con questa ennesima rivoluzione scientifica. Dalla rivoluzione copernicana alla relatività generale, nessun mutamento di paradigma ha messo in causa la sopravvivenza della specie, fino alla scissione dell'atomo e agli arsenali termonucleari delle superpotenze. Finora dal male è venuto il rimedio, finora la specie è sopravvissuta alla minaccia. Toni apocalittici circolano anche nel caso della rivoluzione informatica.

Si potrà evitare l'*info-calisse*, se dalla prevenzione dei mali verrà ancora una volta il rimedio».

5. COME RICONCILIARE INTELLIGENZA ARTIFICIALE E SCIENZA DEI PRINCIPI PRIMI: QUARTERONI

«Generalizzando, e stimolando la nostra immaginazione, penso che oggi si stia sperimentando un cambio di paradigma. Negli studi di ingegneria o delle scienze STEM, si è abituati a pensare che il nostro ruolo sia quello del *problem solver*. Io penso che questo ruolo sarà sempre più appannaggio dell'intelligenza artificiale, con gli strumenti che essa sta sviluppando.

A noi rimane un ruolo fondamentale, unico e non delegabile, quello del *problem setting*. Dobbiamo sapere *come porre il problema*. I problemi sono molto complessi. Questa è la vera frontiera: il *problem setting* richiede conoscenza di dominio, quindi conoscenza della teoria, quella, diciamo, codificata nelle leggi fondamentali della natura e quindi nei modelli matematici che stabiliscono quali debbano essere i dati e le variabili rilevanti, e quali le relazioni fra di essi.

Non è vero, secondo me, che l'Intelligenza Artificiale sancirà la fine della scienza. Io sperimento in quello che faccio ogni giorno quanto ci sia ancora da capire e da scoprire. Più si avanza, più ci si rende conto della nostra ignoranza. Perché più si fa, più si scopre e più si capisce quanto sia l'universo della conoscenza. Il perimetro della conoscenza umana non ha un limite superiore.

Allora, da questo punto di vista, credo che ci sia un grandissimo spazio ancora per tutti noi. E in particolare per la ricerca fondamentale, per poter aiutarci a spostare sempre più in là i limiti della nostra conoscenza. E in definitiva, il progresso della nostra umanità».

6. INTELLIGENZA ARTIFICIALE, CONDOTTA UMANA E TEORIA: UNA PROSPETTIVA DALL'ANALISI ECONOMICA: SCAZZIERI

«La relazione fra strutture e complessità è di importanza centrale per chiarire il ruolo della teoria nell'analisi economica [...] Il concetto di *complessità strutturale* fornisce un'utile chiave di lettura a questo riguardo. Si tratta dell'idea che nei sistemi economici la complessità sia organizzata su più livelli, con caratteristiche di *ridondanza* che rendono possibile una pluralità di distinte rappresentazioni della loro struttura [...] Un esempio illuminante è la possibilità di descrivere le interdipendenze fra attività produttive in un sistema complesso di divisione del lavoro facendo riferimento, in alternativa, a una

rappresentazione “orizzontale” della struttura in termini di flusso circolare [...] oppure ad una rappresentazione “verticale” della struttura in termini di una collezione di sub-sistemi verticalmente integrati costruiti a partire dai beni finali di consumo e di investimento [...] oppure a partire da particolari risorse non prodotte [...] Ridondanza e pluralità di rappresentazioni strutturali richiamano l'attenzione sui criteri di visualizzazione delle interdipendenze [...] e sul ruolo degli spazi concettuali come generatori di *rappresentazioni virtuali* che permettano all'immaginazione scientifica di andare oltre distinzioni consolidate per mettere a fuoco le forme di interdipendenza che via via si modificano lungo le traiettorie di cambiamento strutturale [...].

Possiamo aggiungere, anche alla luce di quanto si è detto in precedenza, che lo stesso lavoro teorico trova, in ultima analisi, un significativo fondamento proprio nell'apertura alla *molteplicità di teorie*, dal momento che questa molteplicità esprime concretamente la *pluralità di possibili rappresentazioni* di una struttura di interdipendenze e l'apertura della teoria ai cambiamenti strutturali, a volte epocali, che hanno luogo nella storia».

7. GUARDARE AL PASSATO PER CAPIRE IL PRESENTE E GOVERNARE IL FUTURO

Le analisi e le valutazioni di Colleghi così autorevoli mostrano nel seguito come da diversi punti di vista la razionalità scientifica e la dimensione storica non temono i cambiamenti radicali, nella consapevolezza che l'Umanità può affrontare il futuro sapendo che nello stesso si presentano spesso fattori ignoti. Per questo è sempre necessario avere uno spirito costruttivamente critico dal quale possono giungere correttivi rispetto a quello che talvolta appare un progresso che si autoalimenta.

ANGELO MARIA PETRONI^(*)

LA SCOPERTA SCIENTIFICA E LA SUA LOGICA NELL'ERA DELL'INTELLIGENZA ARTIFICIALE

LA DEFINIZIONE DI "SCIENZA PURA"

Vi è una differenza fondamentale tra ricerca scientifica ed applicazioni tecnologiche, le quali possono rispondere ai comandi che derivano dalla legge dell'offerta e della domanda – in economie di mercato – o di un potere politico – nei sistemi a vario titolo dirigisti.

La linea di confine tra ricerca pura e ricerca finalizzata, e tra ricerca pura e ricerca tecnologica, per quanto possa rivelarsi in molti casi difficile da tracciare, non è per questo logicamente ed empiricamente meno reale. Le definizioni standard sono date nel *Manuale di Frascati*, nato nel 1963 in ambito OCSE e che è il riferimento internazionalmente accettato. «**Basic research** is experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundations of phenomena and observable facts, without any particular application or use in view». «**Applied research** is original investigation undertaken in order to acquire new knowledge. It is, however, directed primarily towards a specific, practical aim or objective»⁽¹⁾.

Il Manuale di Frascati usa l'espressione diventata più comune, «Basic research». Mentre l'estensione delle due espressioni «pure science» e «basic research» è eguale, l'intensione dei due concetti è diversa. La seconda espressione infatti denota una attività che non è fine a se stessa, ma è funzionale ad altro, ovvero ad attività applicative che sarebbero uno scopo superiore.

La questione ha assunto una importanza cruciale a partire dalle ricerche nucleari negli Stati Uniti durante la seconda guerra mondiale, ed alla pro-

(*) Sapienza Università di Roma.

(1) OECD 2015, capitolo 2: 2.25; 2.29. Interessante è anche la categoria di *Experimental development*: 2.32 «Experimental development is systematic work, drawing on knowledge gained from research and practical experience and producing additional knowledge, which is directed to producing new products or processes or to improving existing products or processes».

gressiva implementazione di politiche di finanziamento pubblico alle imprese private per la ricerca applicata/tecnologica, anche militare.

Nel 1947, quando la “Big Science” stava prepotentemente emergendo – appunto in connessione con le ricerche di fisica nucleare fortemente spinte dalle finalità belliche – Edward Shils così pose la questione delle ragioni della scienza pura: «(1) a clear and fundamental distinction between pure and applied science, and (2) the conviction that the knowledge of the laws of Nature is a good in itself».

Shils rifiutava l'idea che la scienza applicata fosse una autonoma attività di scoperta: «applied scientific work has produced very little insight into the laws of Nature»⁽²⁾. Coincidendo con il giudizio di Louis Pasteur dato molti decenni prima, ed in piena epoca positivistica, per il quale non esisteva una autonoma «scienza applicata», ma esistevano solo applicazioni della scienza.

La conseguenza normativa, sul piano pratico, era per Shils che i fondi pubblici per la ricerca non devono orientare la *direzione* della ricerca, ma soltanto determinare la *quantità* della ricerca. Gli scienziati, e solo gli scienziati, devono decidere come utilizzare i fondi. Questa distinzione è essenziale, in una società democratica.

SCIENZA PURA E METODO SCIENTIFICO

Che la scienza moderna, sin dalle sue origini, si definisca e si distingua dalle altre attività conoscitive per un metodo peculiare, è sempre stato un concetto epistemologico ineludibile. A partire dal “metodo galileano”, riassunto nella diade delle «sensate esperienze» e delle «necessarie dimostrazioni», vi è sempre stato un dibattito su questioni come la unicità del metodo per scienze diverse, come la fisica o la chimica, o per scienze nelle quali un elemento costitutivo è la dimensione storica, quali la biologia evoluzionistica; o su questioni come l'esistenza o meno di un rapporto di priorità temporale o di priorità logica tra la ricerca teorica e la ricerca osservativa e sperimentale. Tuttavia la tesi che il metodo scientifico o non esista, o sia sostanzialmente irrilevante, è stata sostenuta soltanto da punti di vista epistemologici del tutto minoritari, come quello di Paul Feyerabend.

Al centro del metodo scientifico, come processo e come finalità, si pone la scoperta delle leggi di natura. Come per il metodo scientifico, vi è da sempre stato un dibattito, sia sulla loro struttura, sia sul loro statuto epistemologico. Le due polarità entro le quali il confronto si è svolto, e si svolge ancora, è tra la visione realistica – per la quale le leggi descrivono la struttura ontologica

(2) Shils 1947, pp. 80-81.

della natura – e la visione *convenzionalistica*, per la quale le leggi di natura hanno o un carattere puramente definitorio, o un carattere di mera classificazione dei dati, senza alcun carattere di unicità – in entrambi i casi rivestendo un carattere puramente sintattico e non semantico.

Tuttavia, come per il metodo scientifico, soltanto da punti vista epistemologici del tutto minoritari è stata affermata o l'impossibilità o l'irrelevanza delle leggi di natura insieme come parte essenziale e come finalità della scienza.

METODO SCIENTIFICO, LEGGI DI NATURA, ED IA

Il confronto epistemologico sul metodo scientifico ed *insieme* sulle leggi di natura è mutato drasticamente negli ultimi anni, in conseguenza dell'avvento e dello sviluppo esponenziale dell'Intelligenza Artificiale. La tesi avanzata è che Intelligenza Artificiale e "Big Data" – oramai quasi una realtà unica, se non un concetto unico – rendono o renderanno inutile l'attività di "Theory building", ovvero della scoperta delle teorie e della costruzione di modelli teorici. Ovvero, uno dei cardini del metodo scientifico.

La tesi è stata sostenuta nel modo più netto da Chris Anderson: con i "*Big-Data*" saremmo a «the end of theory», perché «the data deluge makes the scientific method obsolete». La «new availability of huge amounts of data, along with the statistical tools to crunch these numbers, offers a whole new way of understanding the world. Correlation supersedes causation, and science can advance even without coherent models, unified theories». Di conseguenza «correlation is enough», così che «we can stop looking for models. We can analyze the data without hypotheses about what it might show. We can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find pattern where science cannot». Ad esempio, «Google's founding philosophy is that we don't know why this page is better than that one: If the statistics of incoming links say it is, that's good enough. No semantic or causal analysis is required». Pertanto «there's no reason to cling to our old ways. It's time to ask: What can science learn from Google?». *Sarebbe non meno che la fine del metodo scientifico nato con la Rivoluzione scientifica rinascimentale*⁽³⁾.

La tesi di Anderson è del 2008. Da allora i progressi dell'Intelligenza Artificiale sono stati formidabili. Non si tratta solo di progressi quantitativi. Siamo infatti passati dalla Intelligenza Artificiale "tradizionale", nata negli anni Cinquanta, concepita come implementazione automatica dei principi della

(3) Anderson 2008.

logica – e anche delle regole grammaticali delle lingue – al “Deep Learning”, il quale prescinde largamente da formalismi logici, e mira ad individuare *patterns* nei dati attraverso una pluralità di tecniche statistiche e bayesiane implementate da reti neurali e dalla stratificazione di esse.

LE LEGGI DI NATURA COME “ONLY FINDING PATTERNS IN THE DATA THAT
HAVE BEEN OBSERVED”

Crediamo sia necessario, per comprendere la sfida che il “Deep Learning” pone alla stessa ragion d’essere della scienza pura, ovvero l’attività di “Theory building”, partire da come l’Intelligenza Artificiale “tradizionale” – o, piuttosto, una parte di essa – riteneva di essere in grado di riprodurre la scoperta delle leggi di natura.

In un celebre articolo del 1973 Herbert Simon affermò che «Law-discovery means only finding patterns in data that have been observed. A law-discovery process is a process for recoding, in a parsimonious fashion, sets of empirical data. A normative theory of scientific discovery is a set of criteria for evaluating law-discovery processes»⁽⁴⁾.

Che la definizione di Simon – sulla quale si basa tutto il poderoso lavoro suo e della sua scuola di costruzione di modelli di Intelligenza Artificiale che possano avere la forza logica e l’adeguatezza storiografica sufficiente a “risco-primare” alcune delle grandi leggi naturali della storia della scienza, dalla Terza legge di Keplero alla legge di Coulomb al ciclo di Krebs – sia una definizione generalmente adeguata da un punto di vista descrittivo, è una questione alla quale ho dedicato una parte importante delle mie ricerche, giungendo a una conclusione negativa. Sono profondamente grato alla memoria di Simon per l’attenzione che negli anni ed in diverse sedi egli volle avere nei confronti delle mie posizioni⁽⁵⁾.

Al di là della questione dell’adeguatezza generale, logica e storiografica, della definizione simoniana di scoperta delle leggi scientifiche, certo è che le due grandi scoperte paradigmatiche della fisica contemporanea, le leggi fondamentali della microfisica quantistica, e le leggi fondamentali della macrofisica relativistica, non hanno avuto nulla a che vedere con un «recoding, in a parsimonious fashion, sets of empirical data». *Nulla vi era nei “dati” disponibili a Planck, e ancor meno vi era nei “dati” disponibili ad Einstein, da cui*

(4) Simon 1973, p. 327.

(5) Cf. Simon 1998. Si vedano anche Petroni 1988, Petroni 1992a, Petroni 1992b, Petroni 1997, Petroni 1998. Il carattere strettamente induttivo del machine learning è molto bene evidenziato da Ippoliti 2023.

si potesse derivare, per interpolazione, estrapolazione, o con un qualsiasi altro ragionamento induttivo-statistico, che la materia e l'energia avessero una natura discreta, che la massa fosse una quantità vettoriale e non scalare, e che la luce fosse soggetta all'attrazione gravitazionale.

In questo contesto basterà ricordare quest'ultima nozione. Conseguenza diretta della Relatività generale, la curvatura della luce (solare) venne osservata dall'isola di Principe, sulla costa occidentale dell'Africa, da Arthur Eddington nel 1919, servendosi dell'eclisse di sole avvenuta il 29 maggio. Ma nessuna osservazione e nessun esperimento di ottica, né classica né elettromagnetica, aveva mai registrato un percorso della luce che non fosse rettilineo. Né nei laboratori, né nelle osservazioni dell'universo, anche con strumenti precisi e potenti come erano quelli già allora a disposizione dei fisici e degli astronomi. Epistemologicamente, la Relatività generale introduce un concetto che per *intensione* e per *estensione* non era contenuto nei dati precedenti, *neanche come singolarità*.

Vi è un altro argomento fondamentale a sostegno della tesi appena illustrata. Esso è un argomento che ha una forza logica – nel senso tecnico del termine – minore, ma una estensione descrittivo/storiografica senz'altro maggiore. Si consideri il caso delle leggi di Newton, così come esse vennero formulate nei Principia. Queste leggi, quando venivano congiunte ai dati disponibili all'epoca, permettevano di fare delle predizioni con un certo grado di precisione (questo era particolarmente rilevante per le predizioni astronomiche). Man mano che i dati osservativi e sperimentali diventarono più precisi (ad esempio, con il miglioramento del potere risolutivo dei telescopi) le predizioni diventarono più precise, *senza che vi fosse stato alcun cambiamento nelle leggi newtoniane*. Dal punto di vista logico, spiegazioni e predizioni – almeno per quanto riguarda le leggi deterministiche – sono il risultato di un modello deduttivo. Il più celebre di essi è il cosiddetto modello nomologico/deduttivo, proposto da Carl Gustav Hempel, da Karl Raimund Popper, e da Paul Oppenheim⁽⁶⁾.

Se le leggi di Newton fossero state una derivazione logica dai dati osservativi dei quali egli disponeva, il loro contenuto logico ed empirico non avrebbe, per definizione, potuto eccedere quello dei dati medesimi. Quindi esse non avrebbero potuto permettere, quando congiunte con dati osservativi di precisione maggiore, di dedurre descrizioni di un grado di precisione maggiore di quello dei dati disponibili all'epoca della loro scoperta. Questa tesi può venire applicata a molte delle leggi fondamentali della fisica, come le leggi di Coulomb, la legge di Gauss, o le equazioni di Maxwell.

Nella ovvia diversità dei contesti storici, delle conoscenze teoriche, e degli strumenti matematici, è possibile quindi argomentare che non vi siano diffe-

(6) Per una chiara sintesi espositiva dei modelli di spiegazione Hempel/Popper/Oppenheim si veda Woodward 2014.

renze strutturali, nel rapporto tra esperienza e teoria, fra la rivoluzione nella fisica del XX secolo e la Rivoluzione scientifica del XVI e del XVII secolo.

Valga solo un punto. La Rivoluzione scientifica del XVI e XVII secolo aveva due fondamenti. In primo luogo, l'eliostaticismo/eliocentrismo di Copernico. In secondo luogo, il principio di inerzia di Galileo. Quest'ultimo permette la formulazione delle leggi del moto, e la nascita della meccanica moderna, con l'unificazione della fisica terrestre e della fisica celeste alla quale perverrà Newton.

Nessuna osservazione astronomica e nessun esperimento aveva mai registrato né la rotazione della terra, né la sua traslazione. Per la prima bisognerà attendere alcune esperienze, non concludenti, alla fine del Settecento e, definitivamente, Foucault nel 1851. Per la seconda bisognerà attendere la misurazione della parallasse stellare da parte di Friedrich Wilhelm Bessel nel 1838.

Per quanto riguarda il principio di inerzia, nessuna osservazione ed esperienza, da quelle del senso comune a quelle elementarmente sperimentali dell'era di Galileo – o, se si preferisce, dei calcolatori del Merton College di Oxford nel XIV secolo – mai avevano registrato un moto che permanesse nel medesimo stato se su di esso non agivano forze esterne.

La Rivoluzione scientifica, da Copernico a Newton, passando per Galileo e Keplero, fu resa possibile dal fatto che ogni concetto fondamentale venne introdotto in modo *controinduttivo*. Le osservazioni e gli esperimenti avevano la funzione di controllo della adeguatezza delle ipotesi teoriche – o, più spesso, della adeguatezza dei modelli matematici di predizione dei fenomeni, basati sulla congiunzione tra dati e relazioni legiformi – ma non furono *e non potevano essere* l'origine, né logica né psicologica né storica, dei concetti fondamentali della nuova scienza⁽⁷⁾.

Se le argomentazioni sopra esposte sono corrette, allora si può concluderne che «The end of theory» affermata da Anderson è errata. Nessun «deluge of data», unito ad algoritmi di interpolazione ed estrapolazione, di qualsiasi potenza si assumano, può produrre la scoperta di leggi comparabili a quelle che hanno definito la storia della scienza, e segnatamente della fisica.

Soprattutto, senza la scoperta di queste leggi non vi sarebbe stato affatto il progresso della scienza come esso è stato⁽⁸⁾.

(7) Cf. Petroni 1999.

(8) Cf. Petroni 2020.

“DEEP LEARNING” E SCOPERTA SCIENTIFICA

Se Anderson riteneva che con l'Intelligenza Artificiale «science can advance even without coherent models, unified theories», negli ultimi anni di contro vi sono stati diversi tentativi di dimostrare che modelli di “Deep Learning” sono in grado di riprodurre la scoperta di leggi fondamentali della fisica partendo esclusivamente da dati osservativi e sperimentali.

Oggetto ne sono state, tra le altre, le leggi di Keplero, le leggi di Newton, l'eliocentrismo di Copernico. È stato persino proposto un programma che esamina dei dati consistenti esclusivamente in dei video, e ricerca il minimo insieme di variabili fondamentali che descrivono adeguatamente la dinamica di alcuni sistemi fisici (quali un pendolo doppio, o una fiamma) con la conclusione che «Without any prior knowledge of the underlying physics, our algorithm discovers the intrinsic dimension of the observed dynamics and identifies candidate sets of state variables». L'entusiasmo di alcuni commentatori, per i quali «Artificial Intelligence Discovers Alternative Physics», non è certo unanimemente condiviso⁽⁹⁾.

In generale, il programma di ricerca che mira alla ricostruzione di alcune leggi fondamentali della fisica attraverso il “Deep Learning” è senz'altro di grande interesse. L'interesse fondamentale, ovviamente, non è quello di avere un nuovo strumento per la storiografia della scienza, ma è quello di avere

(9) Chen *et alii* 2021. Cf. Contucci 2024: «In campo scientifico l'utilità dell'AI come strumento è enorme, ma non esiste ancora un'intelligenza artificiale che possa scoprire teorie fisiche a partire da dati sperimentali, né una che possa proporre frammenti di teorie, i cosiddetti modelli. Le notizie sensazionalistiche che circolano a riguardo sono mistificazioni di piccoli risultati mal compresi. Nel campo della matematica invece l'AI moderna ha di recente mostrato la capacità di proporre raffinate congetture che sono poi state dimostrate dagli umani». Come sostiene ancora Contucci, l'Intelligenza Artificiale è al servizio della scoperta scientifica in modi molto concreti. Uno di essi è quello delle scale molecolari. Mentre sappiamo trattare l'equazione di Schroedinger per atomi semplici, come quello di idrogeno, se alcuni atomi si mettono insieme (molecola) risolvere quell'equazione diventa molto complesso se non impossibile. Il problema, dal punto di vista matematico, che emerge nello studio analitico della molecola fatta da n atomi è molto più complesso del corrispondente problema degli n corpi della meccanica classica. I fisici-matematici si esprimono dicendo che non c'è una chimica-matematica. Questo non impedisce di fare ricerca. C'è il settore classico delle simulazioni numeriche nella chimica che ha prodotto, negli ultimi 50 anni, risultati approssimati di altissimo valore, e ora c'è il “Deep Learning” che sta facendo fare progressi molto rilevanti a campi come il ripiegamento delle proteine e la chimica molecolare, bypassando le soluzioni analitiche con il metodo che gli è congeniale, cioè il calcolo delle correlazioni. Queste spesso sono sufficienti per progettare oggetti come nuovi componenti chimicamente attivi, che vengono poi testati in farmaceutica o usati per altre applicazioni (Comunicazione personale).

uno strumento teorico al quale non sarebbe in alcun modo preclusa la possibilità di fare la scoperta di nuove leggi scientifiche. *Con buona pace di Anderson, l'IA non porterebbe alla "End of theory", ma sarebbe uno strumento per far avanzare la conoscenza teorica.*

Crediamo non sia possibile esprimere un giudizio generale su questo programma di ricerca, per la duplice ragione che i singoli lavori sono estremamente differenziati dal punto di vista metodologico, e i risultati sono anch'essi interpretabili in modo molto differenziato.

Certo gli è che almeno alcuni di essi, valutati da un punto di vista metodologico non separato da una prospettiva storiografica (quest'ultima del tutto necessaria, visto che si tratta di ri-scoperte), presentano profili di incongruità. Ci limitiamo a due esempi.

Il primo riguarda la Terza legge di Keplero: «We considered the problems of deriving Keplers's third law of planetary motion, providing reasoning-based measures to analyze the quality and generalizability of the generated formulae. Extracting this law from experimental data is challenging, expecially when the masses involved are of very different magnitudes. This is the case for solar system, where the solar mass is much larger than the planetary masses. The reasoning module helps in choosing between different candidate formulae and identifying the one that generalize well: using our data and theory integration we were able to re-discover Kepler's third law»⁽¹⁰⁾. La Terza legge di Keplero è una legge puramente cinematica, non dinamica. Il concetto di massa entrerà nelle leggi astronomiche soltanto con Newton. Ed è ovvio che se si considera la Terza legge di Keplero come un qualcosa da ricavare da dati relativi alle masse il procedimento ed il risultato sono del tutto incongrui. Peraltro, la Terza legge di Keplero è una delle poche che, per la sua semplicità, era stata riscoperta dai modelli di Simon: i quali non sono stati in grado di riscoprire né la prima né la seconda legge di Keplero, il cui carattere di universalità è nettamente maggiore.

Il secondo riguarda la legge della gravitazione universale di Newton: «We (re-)discovered Newton's formula for gravitational force from observed trajectories of the Sun, planets, and moons of our solar system, and made accurate estimates of the hidden properties». Gli autori correttamente specificano che «While our method allows us to re-discover Newton's formula and the masses, it is important to note that this was only possible through the use of inductive biases, particularly Newton's second and third law, and spherical symmetry». Ma la specificazione non è sufficiente: assunte la seconda e la terza legge di Newton, e assunto che debba esserci una "hidden property", che ne

(10) Cf. Cornelio *et alii* 2023. Vedasi anche Petroni 1990.

segua la legge della gravitazione universale è difficilmente considerabile come una descrizione di una scoperta di una legge a partire dai dati⁽¹¹⁾.

PROGRESSO AD FINITUM O AD INFINITUM?

La tesi per la quale l'Intelligenza Artificiale dovrebbe portare a modificare, se non ad abolire come vorrebbe Anderson, lo stesso metodo della scienza pura, crediamo possa trarre un elemento di valutazione importante da una delle più classiche questioni epistemologiche, ovvero se il progresso scientifico sia ad finitum o ad infinitum.

Uno dei più grandi fisici del Novecento, Richard Feynman, riteneva che il progresso della scienza sia ad finitum. Come egli scrisse nel 1964, «We are very lucky to live in an age in which we are still making discoveries. It is like the discovery of America – you only discover it once. The age in which we live is the age in which we are discovering the fundamental laws of nature, and that day will never come again. It is very exciting, it is marvellous, but this excitement will have to go. Of course in the future there will be other interests. There will be the interest of the connection of one level of phenomena to another – phenomena in biology and so on, or, if you are talking about exploration, exploring other planets, but there will not still be the same things that we are doing now». Nel futuro egli prevedeva che «There will be a degeneration of ideas, just like the degeneration that great explorers feel is occurring when tourists begin moving in on a territory»⁽¹²⁾.

La visione di Feynman è radicalmente opposta a quella di Karl Popper. Per Popper la ricerca non ha mai fine, ed il progresso scientifico è ad infinitum. Nessuna verità stabilita ad un certo tempo sopravviverà al progredire della ricerca. Il percorso della scienza non è da problemi a soluzioni, ma da problemi più semplici a problemi più complessi⁽¹³⁾.

(11) Lemos *et alii* 2022. Si noti come gli autori affermino anche che «The key assumptions that were required were translational and rotational equivariance, and Newton's second and third laws of motion. Our approach correctly discovered the form of the symbolic force law». In realtà di assunti ve ne sono diversi altri: c'è l'invarianza per traslazione temporale, oltre che quella di traslazione spaziale e rotazionale. Queste tre cose sono la conservazione dell'energia, della quantità di moto e del momento angolare. Si deve inoltre notare come gli autori non inferiscano una legge dell'inverso del quadrato, ma possano solo affermare che essa è la migliore delle altre che hanno testato. Che questo schema di risultato possa essere un *explicatum* tanto logicamente quanto empiricamente adeguato del concetto di scoperta scientifica appare del tutto problematico.

(12) Feynman 1965, p.172.

(13) Popper 1974. La pubblicazione separata della sua autobiografia verrà significativamente intitolata da Popper *Unended Quest*.

La tesi che la ricerca non abbia mai fine può avere due principali presupposti. Il primo presupposto è che la realtà sia di fatto inesauribile. Può non essere infinita, ma è comunque indefinita. Pertanto, dal punto di vista pragmatico, non sono concepibili limiti alla scienza nel secondo significato sopra indicato. Il secondo presupposto è che ogni avanzamento della scienza deriva dall'interazione tra la realtà e le categorie psicologico-concettuali degli scienziati, ad ogni generazione di essi. E deriva anche dai progressi negli strumenti osservativi e sperimentali, che permettono di esplorare livelli sempre più profondi della realtà. Per usare l'immagine di Feynman, l'esplorazione di un territorio non può mai essere limitata ad una dimensione. Di conseguenza il rapporto, in termini filosofici classici, soggetto/oggetto rende la ricerca scientifica essenzialmente senza fine, e sempre progressiva. La visione epistemologica di Federigo Enriques è su questo tema di particolare rilevanza⁽¹⁴⁾.

Una questione correlata, se non derivata, riguarda la dinamica della ricerca scientifica. Un elemento essenziale nella concettualizzazione della dinamica della scienza è il ruolo delle rivoluzioni scientifiche. Dalla nascita della scienza moderna sino agli anni Trenta del Novecento per la fisica (con la rivoluzione relativistica e la rivoluzione quantistica) e sino agli anni Cinquanta del Novecento per la biologia (con la scoperta del DNA), non vi è stato alcun cambiamento "rivoluzionario". Kuhnianamente, non vi è stato alcun "cambiamento di paradigma" da più di un secolo.

Dal punto di vista dell'epistemologia di Popper, questa situazione non ha alcuna rilevanza epistemologica. Il successo passato e presente delle teorie scientifiche non aumenta in alcun modo la loro probabilità di verità. Questa tesi, che è stata anche denominata "The pessimistic induction", per cui ogni teoria della storia della scienza che ha avuto successo per un certo tempo è stata poi confutata, si confronta con i risultati empirici della disciplina della scientometria. Per i nostri scopi, si può qui ricordare il fatto che qualcosa come il novanta per cento di tutti gli scienziati esistiti nella storia sono oggi viventi. La quantità di ricerca svolta negli ultimi anni è esponenzialmente maggiore di quella svolta nei decenni precedenti. Pertanto, se le teorie fondamentali della fisica e della biologia, come sopra indicate, non hanno ricevuto una confutazione nei loro principi fondamentali, "The pessimistic induction" è una tesi che potrebbe avere valore come strumento storiografico, ma non

(14) Cf. Enriques 1907. La scienza è per Enriques una scoperta continua, dove un ruolo cruciale è svolto dalle *ipotesi*, mai riducibili a semplici «recoding, in a parsimonious fashion, sets of empirical data», per usare l'espressione di Simon sopra ricordata. Ovvero, non riducibili a processi induttivistici. I suoi risultati dipendono dall'interazione tra la realtà esterna e la dimensione psicologica degli scienziati. Il che esclude da un lato sia l'idealismo sia l'empirismo soggettivista, e dall'altro sia il materialismo – si può ricordare come *Materialismo ed empiriocriticismo* di Lenin, con la sua "teoria del riflesso", venne pubblicato nel 1909 – sia il logicismo.

come strumento descrittivo della scienza contemporanea. Come ha scritto Ludwig Fahrbach, «[...] whereas past theories were constantly threatened by scientific revolutions and “paradigm shifts”, our current best theories have experienced big boosts to the empirical success in the last few decades, have been entirely stable in that time, and are therefore almost certainly safe from scientific revolutions and “paradigm shifts”»⁽¹⁵⁾.

Se la tesi è corretta, una conseguenza necessaria è che i progressi scientifici del futuro deriveranno essenzialmente da affinamenti delle teorie fondamentali, dovuti alla disponibilità di più potenti mezzi di calcolo, e soprattutto da nuove evidenze osservative e sperimentali. Ma né gli uni né le altre metteranno in discussione le teorie medesime.

La visione del progresso scientifico come ad finitum va senz'altro a rafforzare l'idea che la ricerca di nuove leggi di natura non avrà, e forse non ha già ora, la stessa importanza che ha avuto dalla Rivoluzione scientifica fino alla metà del Novecento. Il progresso scientifico sarà essenzialmente progresso nella raccolta di dati osservativi e sperimentali. A sua volta, il progresso teorico non sarà più, o non sarà più in gran parte, derivato dalla invenzione di teorie controfattuali, come quelle che sono state al centro delle diverse rivoluzioni scientifiche. Il progresso teorico consisterà nella capacità di estrarre dai dati conseguenze rilevanti per la spiegazione e la previsione dei fenomeni. Per riferirsi ai classici schemi dei modelli nomologico-deduttivi o nomologico-induttivi alla Hempel-Oppenheim-Popper, l'incremento del potere descrittivo e predittivo deriverà dall'incremento del contenuto empirico delle condizioni specifiche che sono parte dei modelli logici medesimi, e dalla maggiore potenza di questi ultimi.

Volendosi riferire alla dicotomia oramai classica, quella tra “scienza rivoluzionaria” e “scienza normale” enunciata da Thomas Kuhn nei primi anni Sessanta dello scorso secolo, il futuro della scienza vedrebbe il prevalere della seconda rispetto alla prima.

Vi sono a nostro avviso pochi dubbi che la visione ad finitum abbia come conseguenza, esplicita od implicita, di attribuire al “Deep Learning” un ruolo fondamentale e crescente nella dinamica della scienza, mentre simmetricamente riduce la centralità, se non l'indispensabilità, del “Theory building”. Evidentemente non si tratta di concepire una contrapposizione epistemologica tra le due attività. Ognuna di esse contribuisce al progresso conoscitivo. Ed è anche vero che vi possono essere delle interazioni fruttuose tra le due.

Un'area di interazione senz'altro interessante tra ricerca teorica e “Deep Learning” è la costruzione di modelli del comportamento dei sistemi. Come è stato bene evidenziato da studiosi come Max Black e Mary Hesse⁽¹⁶⁾, il pro-

(15) Fahrbach 2017; p. 5070.

(16) Cf. Hesse 1963 e Black 1962.

gresso della scienza deve molto alla costruzione di modelli, ed anche all'uso di metafore – ovvero alla traslazione di concetti da un ambito di ricerca ad un altro. Modelli e metafore svolgono un fondamentale ruolo euristico per la formulazione delle teorie. Gli algoritmi del “Deep Learning” sono uno strumento molto potente per la costruzione di modelli in questo senso.

Si deve inoltre sottolineare come l'Intelligenza Artificiale/“Deep Learning” stia diventando uno strumento sempre più usato nella ricerca scientifica. La Commonwealth Scientific and Industrial Research Organisation (CSIRO) ha recentemente calcolato che circa l'11% delle pubblicazioni nel campo delle scienze fisiche e circa il 4% nel campo delle scienze della vita sia “AI-related”⁽¹⁷⁾. La costruzione di modelli, e la progettazione di esperimenti, riguarda eminentemente le scienze biomediche, ed anche le scienze fisiche, come la meccanica quantistica. Si tratta di una funzione importante a vantaggio della scienza pura. Ricordando la definizione dalla quale siamo partiti, per la quale «Basic research is experimental or theoretical work».

Tuttavia sul piano pragmatico sembra chiaro che si possa presentare una sorta di trade-off, nel momento in cui si viene a considerare la questione cruciale dell'allocazione delle risorse pubbliche.

Va richiamato il fatto che gli investimenti in scienza pura a livello globale non sono affatto in crescita. Lo si può vedere dall'UNESCO Science Report: towards 2030, pubblicato nel 2015 e rivisto nel 2016. Come in esso si rileva, «(The) shift in research priorities is evident in the amount of research funds currently being allocated to applied science [...]. Researchers are investing more than before in turning a discovery in basic research into a commercially viable and sustainable product or technology with a potentially beneficial socio-economic impact». Ed ancora: «the focus of scientific discovery has shifted from basic research to ‘relevant’ or big science»⁽¹⁸⁾.

Più recentemente, «Data show, across several countries a moderate increase in business-performed basic research and a lower growth of basic research in universities and research centers over the last years (OECD, 2017, OECD Science, Technology and Industry Scoreboard 2017)»⁽¹⁹⁾.

D'altro canto, basare la giustificazione della spesa pubblica per la scienza pura sui ritorni tecnologici che essa avrà, non è pienamente soddisfacente. La ragione è che possiamo vedere come molti progressi nella crescita della conoscenza scientifica, come quelli dell'astronomia e quelli della termodinamica, o non hanno avuto alcuna ricaduta tecnologica, o l'hanno avuta a distanza di decenni o di secoli. Dunque, in un orizzonte temporale che non è quello

(17) Cf. CSIRO 2022, e «The Economist» 2023. Per una ottima visione di insieme dell'IA si veda Contucci 2023.

(18) UNESCO 2015, p.55.

(19) Aspen Institute 2002, p.49.

proprio delle decisioni delle politiche pubbliche. Considerazioni diverse riguardano invece la spesa pubblica finalizzata al progresso tecnologico necessario per l'avanzamento della scienza, ovvero per poter condurre nuovi tipi di osservazioni – come in astronomia – o di esperimenti – come nella fisica delle alte energie. Qui la spesa pubblica trova una giustificazione diretta, rispetto al progresso della ricerca non finalizzata.

Parimenti, si può essere tentati dal credere che l'accettazione sociale della scienza, che è la condizione per cui nelle società democratiche possa formarsi un consenso ampio per allocare risorse ingenti ed anche crescenti, sarà la conseguenza necessaria della percezione, che la gran parte dei cittadini senz'altro ha, dell'utilità del progresso tecnologico. Ma cercare il consenso pubblico nei confronti della scienza sulla base delle ricadute tecnologiche significa indebolire tutte quelle branche della ricerca per le quali le ricadute tecnologiche appunto non sono prevedibili o non lo sono in tempi prevedibili.

È indubbio che la principale attrattività del “Deep Learning” deriva dalla promessa di contribuire alla risoluzione di problemi di tipo applicativo. Questo potrebbe contribuire a rafforzare l'effetto di spiazzamento nell'allocazione delle risorse a scapito della scienza pura, nel senso proprio della ricerca di nuove leggi di natura. Un effetto di spiazzamento che potrebbe riguardare anche la ricerca matematica fondamentale.

Un esempio aneddotico ma a nostro avviso significativo – e ancor più significativo perché è tutto interno all'Intelligenza Artificiale – sono le vicende del Turing Award. Istituito nel 1966, esso ha avuto tra i suoi vincitori scienziati come Marvin Minsky, Allen Newell, Herbert Simon, Stephen Cook. Ognuno di essi ha dato contributi fondamentali nel campo della computazione e dell'Intelligenza Artificiale, corrispondendo perfettamente allo spirito dell'opera di Turing. Attualmente il premio – diventato nel frattempo molto rilevante sul piano economico – è finanziato da una delle aziende informatiche più grandi del mondo. Nel 2019 esso è stato attribuito a Patrick Hanrahan e Edwin Catmull, con la motivazione: «for fundamental contributions to 3-D computer graphics, and the revolutionary impact of these techniques on computer-generated imagery (CGI) in filmmaking and other applications». È evidente come le finalità applicative vengano a prevalere su quelle della ricerca fondamentale⁽²⁰⁾.

(20) Cf. la Association for Computing Machinery (ACM) <https://www.acm.org/>.

CONSIDERAZIONE FINALE

Le considerazioni che abbiamo esposto sono del tutto distinte da ogni posizione che sostenga la intrinseca impossibilità di una Intelligenza Artificiale che possa replicare l'intelligenza umana. La posizione più nota in questo è probabilmente quella di John Searle e, come è noto agli storici della filosofia, la questione può essere significativamente fatta risalire a Cartesio e a Leibnitz, con argomenti che ancora oggi hanno un interesse ed una rilevanza del tutto notevoli.

Personalmente riteniamo che il tentativo di *dimostrare* dei limiti ontologici, o anche solo metodologici, alla capacità di una macchina – comunque concepita o effettivamente costruita – di replicare l'intelligenza umana, non abbia maggior fondamento di quanto ne avesse all'epoca (1872), e ne abbia avuto successivamente, *l'Ignorabimus di Emil Du Bois-Reymond, con l'enunciazione dei sette "misteri" destinati per sempre a sfuggire alla spiegazione scientifica*. Per chi condivida una posizione materialistica l'idea di una mente/cervello umana come realtà unica e di principio non riproducibile, che sia da parte della natura o dell'uomo stesso, è errata.

La nostra tesi per cui l'attività di "Theory building", il "nocciolo duro" della scienza pura, non può essere sostituita dall'Intelligenza Artificiale *come noi la conosciamo oggi* non ambisce e non potrebbe ambire ad affermare alcuna impossibilità di principio. Si tratta di una tesi, certo non particolarmente originale, che deriva esclusivamente da considerazioni di tipo logico e metodologico, guidate da una interpretazione di alcuni momenti significativi della storia della scienza.

BIBLIOGRAFIA

- ANDERSON 2008 = ANDERSON C., *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, «Wired Magazine», <https://www.wired.com/2008/06/pb-theory/>, 2008.
- ASPEN INSTITUTE 2002 = ASPEN INSTITUTE 2002 *In Favor of Pure Science: A Report* Pure-Science-Aspen-Institute-2022.pdf (aspeninstitute.org)
- Association for Computing Machinery (ACM) <https://amturing.acm.org/>
- BLACK 1962 = BLACK M., *Models and Metaphors*, Ithaca, N.Y., Cornell University Press, 1962.
- CHEN *et alii* 2021 = CHEN B., HUANG K., RAGHUPATHI S., CHANDRATREYA I., DU Q., LIPSON H., *Discovering State Variables Hidden In Experimental Data* arXiv:2112.10755, 2021.
- CONTUCCI 2023 = CONTUCCI P., *Rivoluzione Intelligenza Artificiale. Sfide rischi opportunità*, Bari, Edizioni Dedalo, 2023.
- CONTUCCI 2024 = CONTUCCI P., *Intelligenza artificiale: cos'è e quali sono le sue applicazioni*, Enciclopedia Treccani, Appendice XI, Roma, Istituto dell'Enciclopedia Italiana, 2024.
- CORNELIO *et alii* 2023 = CORNELIO C., DASH S., AUSTEL V., JOSEPHSON T.R. GONCALVES J., CLARKSON K. L., MEGIDDO N., B. EL KHADIR B., HORESH L., *Combining data and theory for derivable scientific discovery with AI-Descartes*, «Nature Communications» <https://doi.org/10.1038/s41467-023-37236-y>, 2023.
- CSIRO 2022 = CSIRO, *Artificial Intelligence for Science. Adoption trends and future development pathways*, Artificial Intelligence for Science report - CSIRO, 2022.
- ENRIQUES 1907 = ENRIQUES F., *Il valore della scienza*, *Annuario dell'Università di Bologna*, Discorso letto per la solenne inaugurazione dell'anno accademico 1907-1908, Bologna, Tipografia Monti, 1907, pp. 29-55. Ripubblicato in F. Enriques, *Scienza e razionalismo*, Bologna, Nicola Zanichelli Editore 1912, pp. 2-37.
- FAHRBACH 2017 = FAHRBACH L., *Scientific Revolutions and the Explosion of Scientific Evidence*, «Synthese» 194 (2017), pp. 5039-5072.
- FEYNMAN 1965 = FEYNMAN R., *The Character of Physical Laws*, Cambridge (Mass.) and London, MIT Press, 1965.
- HESSE 1963 = HESSE M., *Models and Analogies in Science*, London, Sheed and Ward, 1963.
- IPPOLITI 2023 = IPPOLITI E., *Guida critica alle intelligenze artificiali*, Milano, Egea, 2023.
- LEMONS 2022 = LEMOS P., Jeffrey N., Cranmer M., Ho S., Battaglia P., *Rediscovering orbital mechanics with machine learning* arXiv:2202.02306 pp. 1-11 <https://doi.org/10.48550/arXiv.2202.02306>, 2022.
- Organisation for Economic Cooperation and Development (OECD), Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*, 7th ed. Paris <http://oe.cd/frascati> and <https://www.oecd.org/publications/frascati-manual-2015-9789264239012-en.htm>, 2015.

- PETRONI 1988 = PETRONI A.M., *Sur quelques positions récentes relatives à la logique de la découverte scientifique*, «Les études philosophiques» LXIII, 4 (1988), pp. 454-479.
- PETRONI 1990 = PETRONI A.M., *I modelli, l'invenzione e la conferma. Saggio su Keplero, la rivoluzione copernicana, e la «New Philosophy of Science»*, Milano, Franco Angeli, 1990.
- PETRONI 1992a = PETRONI A.M., *On Some Problems of the Logic of Scientific Discovery*, in W. H. Newton – Smit J. Tianji (eds.), *Popper in China*, London, Routledge, 1992.
- PETRONI 1992b = PETRONI A.M., *Why Have a Heuristic of Scientific Discovery?*, «International Studies in the Philosophy of Science» VI, 1 (1992), pp. 53-55.
- PETRONI 1997 = PETRONI A.M., *La logica e la storia della scoperta scientifica. Una critica di modelli di H.A. Simon*, «Sistemi intelligenti» IX, 2 (1997), pp. 193-204.
- PETRONI 1999 = PETRONI A.M., *Formalizing Discovery, Discovering Realism*, in R. Rossini Favretti, G. Sandri, R. Scazzieri (eds.), *Incommensurability and Translation. Kuhnian Perspectives on Scientific Communication and Theory Change*, Cheltenham, UK, e Northampton, MA, Elgar, 1999.
- PETRONI 2000 = PETRONI A.M., *El concepto de progreso científico, hoy*, «Revista de Occidente» 5 (2000), pp. 5-26.
- POPPER 1974 = POPPER, K.R., *Autobiography of Karl Popper*, in P.A. Schilpp (ed.), *The Philosophy of Karl Popper*, La Salle, Ill., Open Court Publishing, vol. I, pp. 2-184, 1974.
- POLANYI 1951 = POLANYI M., *The Logic of Liberty*, Chicago, The University of Chicago Press, 1951.
- SHILS 1947 = SHILS, E.A., *A Critique of Planning-The Society for Freedom in Science*, «Bulletin of Atomic Scientists» III, 3 (1947), pp. 80-82.
- SIMON 1973 = SIMON, H.A., *Does Scientific Discovery Have a Logic?*, «Philosophy of Science» XL, 4 (1973), pp. 471-480. Ripubblicato in Simon, *Models of Discovery and Other Topics in the Methods of Science*, Boston Studies in the Philosophy of Science, vol. 54, Dordrecht, D. Reidel Publishing Co., 1977, pp. 326-335.
- SIMON 1998 = SIMON, H.A., *La logica della scoperta: una critica delle posizioni di Petroni*, «Sistemi intelligenti» X, 1 (1998), pp. 137-144.
- «The Economist», *How Artificial Intelligence can revolutionise science*, 14 settembre 2023.
- UNESCO Science Report: towards 2030 Paris, 2015 (rivisto 2016) <https://unesdoc.unesco.org/ark:/48223/pf0000235406>
- WOODWARD 2014 = WOODWARD J., *Scientific Explanation, The Stanford Encyclopedia of Philosophy*, ed. by E. N. Zalta <https://plato.stanford.edu/entries/scientific-explanation/>, 2014 (prima edizione: 2003).

PAOLO CASINI^(*)

LA COSCIENZA DI TURING

Ci si può chiedere se i diritti prioritari della ricerca pura, nell'età dell'intelligenza artificiale, siano esposti all'insidia radicale di una ricerca che tenta il sorpasso della stessa coscienza umana grazie a un suo Ersatz, vero o presunto. Semplificando il *Turing test*, se si pone una domanda a un computer, e si dubita se la risposta sia informatica o umana, si può inferirne che la macchina ha una coscienza. Ma quale grado di coscienza, tra la mente di un essere vivente e il suo surrogato informatico? Dopo più di settant'anni l'ambiguità resta, ed è al centro di infinite controversie sulla AI. Ci si interroga sulla simbiosi tra uomo e macchina, sulle macchine superintelligenti, sulle responsabilità dei loro creatori, e così via, con le conseguenze che ne derivano nella pratica. Lo stesso Alan Turing aveva anticipato attorno al 1950 molti temi di controversie oggi correnti. Si chiedeva che cosa accadrebbe se le macchine pensanti fossero in grado un giorno «to outstrip our feeble powers [...] to take control, in the way that is mentioned in Samuel Butler's *Erewhon*»⁽¹⁾. In una conversazione radiofonica tornò a interrogarsi sull'eventualità che in futuro la specie umana possa essere umiliata dalla crescita di menti più avanzate, o per il sorpasso evolutivo di specie oggi inferiori, ipotesi certamente remota di miliardi di anni; oppure per l'avvento di macchine superintelligenti: «This new danger is much closer. If it comes at all it will almost certainly be within the next millennium. It is remote but not astronomically remote, and is certainly something that can give us anxiety»⁽²⁾.

Un'ansia che sembra oggi aver preso corpo nei robot e negli spauracchi diffusi della *science fiction* mediatica, e ha finito per riverberarsi sulla coscienza degli addetti ai lavori. Anche se molti informatici negano o minimizzano i rischi, o li ritengono tanto remoti da non preoccuparci oggi, altri ancora, tra i più autorevoli, li hanno presi seriamente in esame entro prospettive di lungo periodo, come nel progetto 'secolare' della Stanford University, *One Hundred*

(*) Linceo, Sapienza Università di Roma.

(1) A. M. Turing, *Intelligent machinery, a heretical theory*, p. 131; dattiloscritto di Lecture given to '51 society at Manchester: AMT-B-20, in <https://turingarchive.kings.cam.ac.uk>.

(2) A. M. Turing, *Can digital computers think?*, p. 8; dattiloscritto della conversazione BBC Third Programme, 15 May 1951; AMT-B-5, *ib.*

Year Study of Artificial Intelligence⁽³⁾, o nella *Association for the Advancement of AI* - «longitudinal study to study and anticipate how the effects of artificial intelligence will ripple through every aspect of how people work, live and play»⁽⁴⁾.

Lo stato dell'arte esalta piuttosto gli strumenti di ricerca e i successi già acquisiti in ogni campo, dalla fisica alla medicina, alle comunicazioni, all'automazione in ogni settore dell'economia e della cultura o, tra gli effetti più popolari, l'auto senza pilota, la traduzione istantanea, le risposte dell'oracolo TPS. Il verbo del successo nasce da una prospettiva utilitarista e pragmatista, dove tuttavia le questioni più discusse coinvolgono, al di là della ricerca pura, tutti i valori in gioco. Le tecniche algoritmiche che si 'insegnano' ai robots implicano un'alta specializzazione informatica, ma ai robots non si possono chiedere giudizi di valore. Benéfici o maléfici, sono *wertfrei*. Non c'è dubbio che la 'coscienza' del dilemma di Turing sia premorale. Non ha fini salvo il successo, e il dibattito si sposta così su altri piani e si avvolge in dilemmi tutt'altro che epistemologici. Basti pensare a un aspetto a sé stante: gli impieghi bellici, i droni e i missili intelligenti, che hanno reso tanto più micidiale la minaccia nucleare globale. Le ricadute dell'automazione AI nel lavoro umano riempiono la cronaca quotidiana: nelle grandi aziende informatiche sistemi di algoritmi assumono, controllano, licenziano; le manipolazioni delle immagini usurpano e falsificano il lavoro degli attori; la profilazione 'intelligente' è onnipresente nei media, in rete, nei social, e così via. I danni 'moral' ormai ampiamente denunciati vanno ben oltre le tecniche pubblicitarie importune, le *fake news* o la disinformazione. Potrebbero favorire nella massa degli utenti l'avvento di una ebetudine o abulia diffusa. Ai piani bassi della piramide sociale gli operatori sono meri esecutori, progressivamente sostituibili con robots, mentre le decisioni partono dal vertice gestito da una élite ristretta dei programmatori.

Le tecniche di ricerca AI non si propongono di «conquistare ogni tanto un castello o un villaggio di frontiera», per usare una metafora di David Hume, ma «muovono direttamente alla capitale»⁽⁵⁾. La scienza della mente, un tempo affidata alle elucubrazioni e all'introspezione dei filosofi, grazie alle neuroscienze sta trasformandosi in una scienza esatta del cervello. «Padroni di questo centro – diceva ancora Hume – possiamo sperare di ottenere ovunque una facile vittoria»⁽⁶⁾. La vittoria finale sarà dei robots. Traducendo in simboli i messaggi bioelettrici tra centri neuronali cerebrali, gli algoritmi informati-

(3) <https://ai100.stanford.edu>.

(4) <https://aaai.org>.

(5) D. Hume, *A Treatise of human nature, introduction*; trad. it. *Trattato sulla natura umana*, in *Opere*, a cura di E. Lecaldano e E. Mistretta, Bari 1971, 2 vol., I, p. 7.

(6) *Ib.*

ci leggono progressivamente in tempo reale le interazioni tra rete neuronale e coscienza. Ben al di là dei successi ottenuti in astrofisica, nella fisica delle particelle elementari, in biochimica, in genetica, in medicina, il sorpasso delle macchine superintelligenti sull'intelligenza umana cancellerà i confini incerti tra le scienze dure e le scienze umane e risolverà gli ultimi enigmi del *sapiens*. Il traguardo dell'evoluzione bio-informatica sarà forse una vittoria di Pirro per i programmatori umani, ma schiuderà certo il santo Graal dell'onniscienza.

Ci si può chiedere d'altra parte come l'AI, strumento di conoscenza che opera per partenogenesi, senza organi né istinti vitali, né i cinque sensi, ed è sottratto alla lotta per la vita e alla selezione naturale, potrà tradurre in simboli ed emulare l'infinita gamma di sensazioni, emozioni, bisogni, propri della condizione umana. Considerando le origini stesse della coscienza, ragion d'essere dei viventi sulla terra, si può lasciare alla bacchetta magica della *science fiction* l'artificio di aggiungere sentimenti ed emozioni ai robots. L'intelligenza negli animali e la coscienza nell'uomo non sono forse esiti dell'adattamento all'ambiente, emersi dai milioni di secoli dell'evoluzione naturale? Come saltare tutti le tappe spaziotemporali intermedie e crearne surrogati artificiali disincarnati?

Esperti e profani, com'è ovvio, non escono dal dilemma di Turing quando si chiedono se il successo dell'onniscienza digitale andrà a favore dell'umanità stessa, oppure dei supercomputers, istruiti dall'uomo ma ormai resi autonomi da ogni sudditanza agli umani. Tra varie utopie e distopie si profilano due casi estremi. Da un lato un'epoca di robot servizievoli: un nuovo paradiso terrestre, dove sarà eliminato ogni ostacolo alla fatica pratica della vita sociale, al benessere, ai piaceri, un nirvana ben più rassicurante dell'utopia ormai datata di Huxley, *Brave new World*; dall'altro estremo la *fictio* diabolica di un computer killer, in grado di procedere allo sterminio dei programmatori. Sono reminiscenze di archetipi emergenti dal subconscio della cultura occidentale: il dono della conoscenza nell'Eden e il frutto proibito della tentazione, la punizione dell'ῥβπις di Prometeo, di Dedalo inventore del labirinto e di Icaro sulla macchina volante.

L'opinione sobria valuta i pro e i contro della rivoluzione in atto, sceverando razionalmente ciò che è possibile, probabile, impossibile. Essendo impensabile rinunciare alle grandi aspettative, o fermare progetti resi ormai incontrollabili e sostenuti da enormi interessi, è urgente piuttosto ragionare sulla compatibilità con l'umano: *human compatible*. Stuart Russell, oxoniano e docente di Computer Science a Berkeley, nel suo libro così intitolato, ha discusso un'ampia gamma di varianti: opinioni, strumenti per la prevenzione e il controllo di ciò che risulterebbe incompatibile con l'umano, anche

se tracciare un confine netto è difficile⁽⁷⁾. Si pensi, ad es., alla possibilità di memorizzare e rendere consultabile con un click l'intero patrimonio di conoscenze acquisite. La meta di un archivio universale sembra ormai prossima, ma gli esiti dell'onniscienza archiviata potrebbero risultare anch'essi ambigui. In migliaia di secoli di evoluzione dall'*homo erectus* al *faber* al *sapiens* milioni di generazioni si sono adattate all'ambiente e lo hanno modificato memorizzando e modificando di generazione in generazione patrimoni teorici e pratici di cognizioni. La fatica spesa nell'apprendimento di miliardi di umani per bilioni di ore non è da escludere dal bilancio finale. La reazione al vecchio, la curiosità del nuovo hanno agito da stimolo nelle varie rivoluzioni scientifiche. Ci si può chiedere se il feed-back dalla superintelligenza ai singoli utenti umani potrà stimolare ulteriormente questo processo, oppure favorire abulia e inappetenza del sapere. Se le macchine diventassero protagoniste della ricerca e controllassero le menti umane, potrebbero considerarci come una specie rimasta indietro nell'evoluzione, lasciarci vivere con il loro beneplacito – secondo il “Gorilla problem”, un luogo comune nella letteratura sulla AI – come gli umani conservano la specie protetta del Gorilla nelle foreste⁽⁸⁾.

Ci si può chiedere infine come convivere con questa ennesima rivoluzione scientifica. Dalla rivoluzione copernicana alla relatività generale, nessun mutamento di paradigma ha messo in causa la sopravvivenza della specie, fino alla scissione dell'atomo e agli arsenali termonucleari delle superpotenze. Finora dal male è venuto il rimedio, finora la specie è sopravvissuta alla minaccia. Toni apocalittici circolano anche nel caso della rivoluzione informatica. Si potrà evitare l'*info-calisse*, se dalla prevenzione dei mali verrà ancora una volta il rimedio.

(7) Stuart Russell, *Human compatible. AI and the problem of control*, London 2020. Russell è autore con P. Norvig del manuale istituzionale *Artificial intelligence, A modern approach*, Prentice Hall, 1995, 2003 e successive ristampe.

(8) S. Russell, *Human compatible*, cit., pp. 132 sgg.

ALFIO QUARTERONI^(*)

COME RICONCILIARE INTELLIGENZA ARTIFICIALE E SCIENZA DEI PRINCIPI PRIMI

UNA PREMessa STORICA

Non sono un filosofo della scienza, sono un matematico, uso strumenti di intelligenza artificiale e, ovviamente, cerco di farlo in maniera consapevole (talvolta anche originale).

Vorrei iniziare ricordando che il termine *Intelligenza Artificiale* venne coniato nel 1955, da un gruppo di ricercatori fra i quali John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon con la “proposta di Dartmouth”, ovvero di una conferenza estiva, in cui si proponevano di discutere di temi quali le reti neurali, la teoria della computabilità, l’elaborazione del linguaggio naturale. Fu proprio durante tale conferenza, tenutasi nell’estate del 1956 nel campus del Dartmouth College (NH), che venne presentato il primo programma esplicitamente progettato per abilitare i computer a imitare le capacità di *problem solving* degli esseri umani. La primissima definizione, attribuita allo stesso McCarthy, recita «the science and engineering of making intelligent machines». È grazie a tale definizione che John McCarthy, allora ventottenne, un computer scientist e scienziato cognitivo del MIT (poi dal ’62 passato a Stanford), è oggi riconosciuto come il padre dell’Intelligenza Artificiale (si veda McCarthy 1959).

Un altro precursore dell’IA fu l’inglese Alan Turing. Nel suo celebre articolo *Computing machinery and intelligence*, sulla rivista «Mind» nel 1950 (Turing 1950), egli pose la domanda cruciale sulla possibilità che una macchina potesse pensare. Questa domanda implicava la necessità di una definizione chiara di ‘macchina’ e ‘pensiero’. Turing diede vita al celeberrimo “Test di Turing”, un criterio per valutare se una macchina possa fornire a un interlocutore umano risposte indistinguibili da quelle di un essere umano.

(*) Linceo, Politecnico di Milano; EPFL - Lausanne.

Le origini dell'IA, quindi, affondano le radici in una congiunzione di idee e sforzi visionari, alimentati da riflessioni pionieristiche sulla possibilità di conferire intelligenza alle macchine. Da queste origini, l'IA ha attraversato un percorso di sviluppo accelerato, passando attraverso molteplici fasi di innovazione tecnologica e ridefinendo costantemente i confini della nostra interazione con la tecnologia intelligente. Si veda McCarthy 1989, 1990, 1996.

IL DATA DELUGE

Nonostante la sua storia ormai quasi settantennale, l'Intelligenza Artificiale ha tuttavia conosciuto uno sviluppo esponenziale soltanto negli ultimissimi anni, grazie sostanzialmente a tre fattori fondamentali. Il primo è quello dei cosiddetti *big data*. Il secondo è da ascrivere alla disponibilità di computer, o meglio, super computer, specialmente nel cloud, a costi relativamente ragionevoli, che permettono quindi di processare i big data. Il terzo, fondamentale, quello degli algoritmi. Per estrarre conoscenza e modelli dai big data abbiamo bisogno di algoritmi efficienti ed accurati (ad esempio, quelli che consentono di addestrare le reti neurali artificiali) da implementare su questi supercomputer.

Abbiamo incominciato a 'contare' i big data nell'86. O meglio, nell'86 per la prima volta qualcuno si è posto il problema di quanti dati ci fossero nell'universo, diciamo nelle banche dati. Allora il web non era ancora diffuso a livello planetario. Non è molto importante sapere quanto grande sia questa cifra, però ci sono due elementi che mi sembrano molto interessanti.

Il primo è che questo numero di dati raddoppia in media ogni anno e mezzo. Prendendoci qualche libertà interpretativa, potremmo dire in modo un po' disinvolto che ogni anno e mezzo raddoppia la conoscenza dell'universo. (Usiamo l'equazione semplicistica secondo la quale tanti dati significano tanta conoscenza – ovviamente, non è esattamente così!). Bastano quindi meno di due anni per generare un patrimonio di conoscenza dimensionalmente equivalente a quella già presente ad ogni determinato istante. Questa dinamica di crescita è davvero prodigiosa!

Il secondo è che si stima che entro il 2025 si raggiunga la bella cifra di 150 zettabyte. Uno zettabyte corrisponde a 10^{21} byte, ovvero mille miliardi di miliardi di byte. 21 come 21-esimo secolo: il secolo che stiamo vivendo verrà pertanto ricordato come il secolo degli zettabyte.

Da questi numeri impressionanti prende spunto la tesi di Chris Anderson (Anderson 2008), dal titolo provocatorio *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*.

La concezione della "fine della teoria" ha suscitato notevole interesse, soprattutto nel mondo del business e in alcuni ambiti accademici come la data

science. Chris Anderson, ex caporedattore di Wired, ha propagato questa idea affermando che l'abbondanza di dati rende obsoleti i modelli teorici (come quelli basati sui principi primi che governano la natura, l'universo, la fisica in generale), sostenendo che la correlazione tra dati sia sufficiente senza necessità di spiegazioni causali. Questo paradigma sfida il metodo scientifico tradizionale, suggerendo che la correlazione sia sostituita alla spiegazione.

Tuttavia, questa prospettiva presenta fallacie. L'idea che i *Big Data* possano cogliere esaustivamente un'area di ricerca ignora che ogni insieme di dati è influenzato dalla piattaforma e dall'ontologia utilizzate, riflettendo un particolare punto di vista anziché una verità assoluta. L'assunzione che i dati possano esistere senza interpretazione umana è ingannevole: le decisioni umane influenzano la raccolta, l'analisi e la presentazione dei dati.

A questo proposito, si possono fare infiniti esempi, mi limiterò a uno solo. Qualche anno fa la città di Boston ha cercato di individuare le strade con buche utilizzando i dati degli accelerometri presenti negli iPhone degli abitanti. Questo approccio sfrutta i sensori di movimento presenti nei cellulari per rilevare le vibrazioni e i salti che avvengono quando un veicolo passa sopra una buca o una sezione stradale deteriorata. Questi dati, raccolti in modo anonimo e aggregato, possono essere utilizzati per mappare e identificare le strade che richiedono interventi di riparazione o manutenzione. È un esempio interessante di come i dati generati dagli smartphone possano essere sfruttati per migliorare l'infrastruttura urbana e risolvere problemi pratici nella città. In realtà, l'interpretazione dei dati ha portato a una scoperta interessante. Si è constatato che nei quartieri residenziali, vi erano stati molte più segnalazioni di dissesto. Il *bias* era che la distribuzione degli iPhone fosse più accentuata nei quartieri residenziali (abitati da persone più abbienti) che in quelli popolari, o poveri.

La conclusione ovvia è che il dato di per sé è portatore di conoscenza, ma può talvolta essere portatore di una falsa conoscenza. Ancora una volta saper trattare i dati richiede moltissima teoria e qui ritorniamo al punto di partenza. La teoria è quella sviluppata dalla conoscenza fondamentale. L'analisi dei dati richiede sempre una cornice interpretativa. Nascondere l'interpretazione dietro l'apparenza di neutralità o oggettività è illusorio. Ma c'è di più: la conoscenza del 'contesto' che ha generato quei dati ha un valore straordinario e irrinunciabile. I modelli matematici servono a 'interpretare' quel contesto, rappresentandolo con equazioni che, in molti casi (nelle scienze naturali, nelle scienze della vita, nelle scienze economiche, nei processi tecnologici), traducono i principi primi della fisica, quelli che affermano verità indiscutibili, quali ad esempio i principi di conservazione della massa, dell'energia, di equilibrio delle forze, le leggi della termodinamica, etc. In altri termini, modelli che codificano un patrimonio di conoscenza straordinario e irripetibile, generato da giganti del passato che rispondono ai nomi di Keplero, Galileo,

Newton, Pascal, Bernoulli, Laplace, Einstein, Maxwell e Schroedinger, giusto per citarne alcuni.

MACHINE LEARNING E ALGORITMI DATA-DRIVEN

Il Machine Learning è una branca dell'IA che si concentra sulla creazione di modelli che possono apprendere dai dati (si veda ad esempio Mitchell 1997). Invece di programmare esplicitamente un algoritmo per eseguire un compito, nel machine learning si forniscono dati di addestramento al sistema e il modello impara autonomamente a eseguire il compito. Quindi, possiamo considerare il machine learning come una tecnologia fondamentale all'interno del campo più ampio dell'IA. Le Reti Neurali Artificiali (ANN), a loro volta, sono un sottosettore specifico del machine learning. Si tratta di modelli ispirati dalla struttura e dalla funzione del cervello umano. Sono composte da unità chiamate neuroni artificiali, organizzati in strati, e queste reti possono apprendere relazioni complesse nei dati attraverso il processo di addestramento. In particolare, il deep learning è una sottocategoria delle reti neurali artificiali che prevede l'uso di reti neurali profonde, composte da molti strati, per apprendere rappresentazioni sempre più astratte dei dati (si vedano ad esempio Goodfellow *et al.* 2016 e Aggarwal 2018).

In estrema sintesi, possiamo dire che l'Intelligenza Artificiale abbraccia il concetto più ampio di creare sistemi intelligenti, mentre il Machine Learning è una tecnologia fondamentale all'interno dell'IA, e le Reti Neurali Artificiali sono un metodo specifico di implementazione del machine learning, particolarmente efficace per problemi complessi e grandi quantità di dati.

Le reti neurali profonde vengono addestrate grazie alla disponibilità di big data che costituiscono il cosiddetto training dataset. Il celebrato GPT (Generative Pre-training Transformer) è un modello linguistico generativo general purpose basato sul *deep learning* (basato sull'architettura di rete neurale Transformer) applicato all'NLP (*Natural Language Processing*) che usa reti neurali allenate da contenuti immessi da grandi fonti e che in modo completamente automatico comprende e genera testo stimando la probabilità che la parola successiva abbia senso nel contesto di riferimento. Inoltre il modello è in grado di effettuare molte elaborazioni testuali dalla traduzione alla sintesi di testi, alla produzione di contenuti redazionali molto lunghi (libri, documenti, sceneggiature, ecc.). GPT-3 è stato addestrato con quasi 800GB di informazioni testuali recuperate dalla rete, da Wikipedia e dai libri digitalizzati. Tutte queste parole, grazie alle reti neurali, sono messe in connessione tra loro (utilizzando 175 miliardi di parametri) e consentono al modello di creare un modello statistico/previsionale del linguaggio.

Questi numeri ci potrebbero indurre a pensare che effettivamente dai dati si possa ricostruire tutto, ma è ancora un'illusione. ChatGPT, al pari di una qualsiasi rete neurale, funziona perché ha soggiacente un modello, la cui costruzione è decisa dal programmatore, quindi dall'uomo, al quale è demandata la scelta dei parametri e dei cosiddetti iperparametri (quelli che caratterizzano la struttura della rete neurale). Dal canto suo, la scelta di parametri e iperparametri è effettuata grazie alla minimizzazione di una funzione obiettivo e gli algoritmi di minimizzazione sono ancora basati su teorie matematiche rigorose (dunque su leggi ed equazioni). Quindi non facciamoci indurre dalla tentazione di pensare che dal dato grezzo si possa estrarre qualunque tipo di conclusione in modo automatico.

MODELLI MATEMATICI GUIDATI DAI PRINCIPI FISICI

Nella sua relazione, il Professor Petroni ha parlato di principi fondamentali della fisica, di leggi di natura, quelle che, fino ai nostri giorni, ci hanno consentito di comprendere la Natura, l'Universo, e la realtà che ci circonda attraverso le equazioni e il formalismo matematico. In effetti, i modelli matematici basati su leggi fisiche sono costruiti sulla comprensione delle leggi e dei principi fondamentali che governano un determinato fenomeno fisico (qui inteso in senso lato), o un sistema di tipo generale. Sono formulati utilizzando equazioni matematiche che descrivono le relazioni causali tra le variabili coinvolte nel sistema e fra quelle e i dati del problema e richiedono una conoscenza approfondita della fisica sottostante. Si veda, ad esempio (Quarteroni 2017, 2020).

I dati alimentano i modelli matematici in termini rigorosi e controllati, fissando condizioni iniziali e al contorno, parametrizzazioni fisiche, termini sorgenti e forzanti. Dati sovrabbondanti (laddove ce ne fossero) non sarebbero opportuni per far funzionare correttamente i modelli matematici. Peraltro, non è nemmeno vero che sempre, in ogni ambito, oggi, vi siano tanti dati a disposizione. A volte abbiamo un problema di scarsità di dati. Come, ad esempio, quando si vogliano costruire algoritmi di intelligenza artificiale per diagnosi precoci. Se al pronto soccorso di un ospedale si presenta un paziente che manifesta (diciamo) un problema di ritmo cardiaco, si vorrebbe costruire algoritmi (di Machine Learning) in grado di decidere in tempo reale se quella persona debba essere ricoverata, oppure tenuta sotto osservazione al pronto soccorso, o ancora rispedita tranquillamente a casa. Quindi emettere codice rosso, giallo, oppure verde. (Ovviamente poi la decisione spetterebbe al medico, gli algoritmi si limiterebbero a fornire "raccomandazioni"). Si tratta di un tipico caso in cui gli algoritmi dell'intelligenza artificiale potrebbero

coadiuvare i medici nella primissima diagnosi. Aiutare, non sostituire, naturalmente. Ma non sempre si disporrà di centinaia o migliaia di pazienti che si sono sottoposti prima a indagini per alimentare, come si dice, il dataset del training dell'addestramento degli algoritmi di machine learning. Questo è un tipico caso di *scarce data*, ovvero in cui ci sono troppo pochi dati per alimentare le reti neurali.

SCIENTIFIC MACHINE LEARNING, OVVERO COME CONIUGARE MODELLI MATEMATICI E ALGORITMI DATA-DRIVEN

Allora, anche se può sembrare paradossale, potrebbe essere un modello matematico a surrogare questa scarsità di dati, generando soluzioni numeriche (grazie ai computer) che simulino quel problema di ritmo cardiaco da utilizzare come dati di training delle reti neurali stesse. Questo modello matematico dovrebbe essere in grado di descrivere correttamente la fisiologia del cuore umano, grazie alle equazioni di Maxwell per la simulazione del campo elettrico, alle equazioni della meccanica dei continui per simulare la deformazione del miocardio (il suo rilassamento e la sua contrazione nelle fasi diastoliche e sistoliche), le equazioni di Navier-Stokes per la dinamica dei fluidi (il sangue, nella fattispecie, che fluisce negli atri e nei ventricoli). Nomi di grandi scienziati che hanno trovato leggi che sono invarianti nel tempo e nello spazio. Bene, queste leggi di natura sono utilizzate per produrre soluzioni che avranno il ruolo di dati per alimentare le reti neurali. Ecco, questo è un primo esempio da cui si comprende come le leggi di natura possono essere messe al servizio delle reti neurali (si veda ad esempio Bucelli *et al.* 2022, Fedele *et al.* 2023).

Vi sono poi tanti altri ambiti in cui le reti neurali si possono porre al servizio dei modelli matematici basati su leggi di natura. Ad esempio, se si scopre un nuovo materiale, probabilmente non conosceremo le sue leggi costitutive, ovvero quelle relazioni che legano, a livello microscopico, gli sforzi e le deformazioni. Questa mancanza di informazione priva il modello matematico di un tassello fondamentale.

Possiamo invece usare le reti neurali basate su dati, naturalmente, per inferire questa relazione, anche se non la legge in maniera formale, perché ancora le reti neurali non riescono a sintetizzare la conoscenza in formule. Quello che fanno è stabilire delle relazioni input-output.

I due esempi precedenti, che fanno capire come la nuova rivoluzione del Machine Learning (che come tutte le rivoluzioni tecnologiche genera inquietudine), vada vista anche come opportunità. È un arricchimento di conoscenza, ha il suo ruolo molto distinto dal ruolo delle leggi di natura. Ma è un momento in cui i due approcci possono trovare un momento di sintesi

straordinariamente potente. Ecco, io penso che questa sia veramente oggi una pista importante da coltivare.

Non avere un punto di vista duale ed alternativo, in cui l'una cosa è contrapposta all'altra, ma capire che le due cose possono effettivamente essere sinergiche.

Gli algoritmi *data-driven* (come quelli di ML e più in particolare le ANN) si basano sull'analisi dei dati per estrarre modelli, tendenze o relazioni. Non presuppongono necessariamente una comprensione approfondita delle leggi fisiche alla base del sistema. Possono essere utilizzati quando le relazioni complesse nei dati non sono facilmente descritte da modelli basati sulle leggi fisiche.

I modelli matematici basati sulle leggi fisiche, gli algoritmi data-driven e gli algoritmi di machine learning rappresentano approcci diversi per affrontare problemi e ottenere previsioni o informazioni utili in contesti specifici.

Come ho precedentemente osservato, in alcuni casi, è possibile integrare modelli basati sulle leggi fisiche con approcci data-driven o di machine learning per migliorare la precisione delle previsioni. Ad esempio, un modello fisico può essere adattato utilizzando dati sperimentali per tener conto di effetti non considerati inizialmente o per meglio identificare parametri e coefficienti che lo caratterizzano. Oppure si può usare per generare dati di training per algoritmi data-driven nei casi in cui questi dati di addestramento siano scarsi (è questo molto spesso il caso dei dati medici generabili da pazienti).

Mancanza di comprensione: nei casi in cui le leggi fisiche non siano ben comprese, gli approcci *data-driven* o di machine learning possono essere preferiti per ottenere modelli utili dai dati disponibili. Ma anche per generare modelli matematici dimensionalmente ridotti, pertanto con costi computazionali nettamente inferiori.

In conclusione, la scelta tra modelli basati sulle leggi fisiche, algoritmi data-driven e algoritmi di machine learning dipende dal contesto specifico, dalla disponibilità di dati e dalla comprensione del fenomeno in esame. In molti casi, l'approccio più efficace è una combinazione di questi metodi.

Ritengo che questo sia oggi l'atteggiamento corretto da tenere. Pensare che questi strumenti potentissimi di intelligenza artificiale possono aiutarci a prendere delle decisioni più informate.

Se abbiamo usato ChatGPT, ci saremo accorti che la qualità della risposta dipende significativamente dalla qualità della domanda che facciamo, attraverso il prompt. Al punto che è sorta recentemente una nuova categoria di data scientists, i cosiddetti *prompt engineers*, che sanno ormai come porre le domande in modo tale da ottenere risposte molto accurate.

Generalizzando, e stimolando la nostra immaginazione, penso che oggi si stia sperimentando un cambio di paradigma. Negli studi di ingegneria o delle scienze STEM, si è abituati a pensare che il nostro ruolo sia quello del *pro-*

blem solver. Io penso che questo ruolo sarà sempre più appannaggio dell'intelligenza artificiale, con gli strumenti che essa sta sviluppando.

A noi rimane un ruolo fondamentale, unico e non delegabile, quello del *problem setting*. Dobbiamo sapere *come porre il problema*. I problemi sono molto complessi. Questa è la vera frontiera: il problem setting richiede conoscenza di dominio, quindi conoscenza della teoria, quella, diciamo, codificata nelle leggi fondamentali della natura e quindi nei modelli matematici che stabiliscono quali debbano essere i dati e le variabili rilevanti, e quali le relazioni fra di essi.

Non è vero, secondo me, che l'Intelligenza Artificiale sancirà la fine della scienza. Io sperimento in quello che faccio ogni giorno, quanto ci sia ancora da capire e da scoprire. Più si avanza, più ci si rende conto della nostra ignoranza. Perché più si fa, più si scopre e più si capisce quanto sia l'universo della conoscenza. Il perimetro della conoscenza umana non ha un limite superiore.

Allora, da questo punto di vista, credo che ci sia un grandissimo spazio ancora per tutti noi. E in particolare per la ricerca fondamentale, per poter aiutarci a spostare sempre più in là i limiti della nostra conoscenza. E in definitiva, il progresso della nostra umanità.

BIBLIOGRAFIA

- McCARTHY 1990 = J. McCARTHY, *Formalizing Common Sense: Papers by John McCarthy*, Ablex Publishing Corporation, 1990.
- MCCARTHY 1989 = J. McCARTHY, *Artificial Intelligence, Logic and Formalizing Common Sense*, in R. Thomason (ed.), *Philosophical Logic and Artificial Intelligence*, Kluwer Academic, 1989.
- MCCARTHY 1996 = J. McCARTHY, *Defending AI research: A collection of essays and reviews*, «CSLI lecture notes» 49, Center for the Study of Language and Information, 1996
- ANDERSON 2008 = C. ANDERSON, *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, «Wired», 23 June 2008, <https://www.wired.com/2008/06/pb-theory/>.
- MITCHELL 1997 = T. MITCHELL, *Machine Learning*, McGraw-Hill, 1997.
- AGGARWAL 2018 = C. C. AGGARWAL, *Neural Networks and Deep Learning. A Textbook*, Springer, 2018.
- GOODFELLOW *et al.* 2016 = Y. GOODFELLOW, Y. BENJO and A. COURVILLE, *Deep Learning*, MIT Pressbook, 2016.
- QUARTERONI 2017 = A. QUARTERONI, *Numerical Models for Differential Problems*, Springer Nature, 2017.
- QUARTERONI 2020 = A. QUARTERONI, *Le equazioni del cuore, della pioggia e delle vele. Modelli matematici per simulare la realtà*, Zanichelli, 2020.
- BUCELLI *et al.* 2022 = M. BUCELLI, A. ZINGARO, P. C. AFRICA, I. FUMAGALLI, L. DEDÈ and A. QUARTERONI, *A mathematical model that integrates cardiac electrophysiology, mechanics, and fluid dynamics: Application to the human left heart*, «Int. J. Num. Meth. Bioeng.», First published: 29 December 2022, <https://doi.org/10.1002/cnm.3678>.
- FEDELE *et al.* 2023 = M. FEDELE, R. PIERSANTI, F. REGAZZONI, M. SALVADOR, P. C. AFRICA, M. BUCELLI, A. ZINGARO, L. DEDÈ and A. QUARTERONI, *A comprehensive and biophysically detailed computational model of the whole human heart electromechanics*, «Computer Methods in Applied Mechanics and Engineering» 410, 15 May 2023, 115983.

ROBERTO SCAZZIERI^(*)

INTELLIGENZA ARTIFICIALE, CONDOTTA UMANA E TEORIA: UNA PROSPETTIVA DI ANALISI ECONOMICA

INTELLIGENZA ARTIFICIALE, CONDOTTA UMANA E REGOLE DI INFERENZA

Il tema dell'intelligenza artificiale è al tempo stesso nuovo e antico. L'idea fondamentale rinvia alla questione della τέχνη, e quindi al tema degli strumenti, dei 'meccanismi', costruiti per facilitare e talora sostituire azioni umane volte a conseguire obiettivi. L'intelligenza artificiale e la tecnologia appartengono alla stessa sfera della condotta umana: l'ambito delle azioni in vista di obiettivi attraverso *azioni indirette* che aumentino l'efficacia delle prime. Un aspetto importante, e anch'esso antico, della costruzione di strumenti è la realizzazione di meccanismi volti ad accrescere l'efficacia del ragionamento umano, soprattutto nel senso di accrescere la capacità di derivare conclusioni da premesse, e quindi di accrescere la capacità di generare conoscenza indiretta. Un aspetto caratteristico di questi meccanismi è il potenziamento di *regole di inferenza*, nel presupposto che attraverso la ricostruzione razionale e riproduzione delle modalità di ragionamento che generano conoscenza indiretta sia possibile costruire strumenti che accrescano l'efficacia di quei ragionamenti. Questa categoria include strumenti che accrescono la capacità di orientarsi nello spazio e nel tempo, come bussola e orologio, così come strumenti volti ad accrescere memoria e capacità di calcolo, dalle "arti della memoria" studiate e praticate nel Medioevo e Rinascimento (Rossi 1960; Yates 1966; Bolzoni 2003) a meccanismi computazionali come gli elaboratori elettronici.

Le premesse dell'intelligenza artificiale appartengono alla sfera della costruzione e utilizzazione di strumenti. A questo proposito vale la pena di sottolineare che gli strumenti riflettono sia gli obiettivi a cui sono rivolti sia le modalità di condotta che gli strumenti mirano a ricostruire e potenziare. Nel

(*) Linceo, Alma Mater Studiorum Università di Bologna; Gonville and Caius College e Clare Hall, Cambridge.

caso di strumenti volti a potenziare la capacità di ottenere conoscenza indiretta, le modalità di condotta rilevanti riguardano (i) la *capacità di visualizzazione* per il conseguimento di obiettivi; (ii) la *capacità di selezione* per mettere a fuoco aspetti significativi nell'ambiente rilevante; (iii) la *capacità di integrazione* fra criteri di visualizzazione diversi quando l'ambiente si modifica nelle sue componenti e/o nella sua struttura (ad esempio quando ne aumenta il grado di complessità).

FRAMING, STRUTTURE, SPAZI CONCETTUALI

Al centro dell'intelligenza artificiale, così come di altri strumenti per accrescere l'efficacia delle procedure per ottenere conoscenza indiretta, è la ricerca di regole di inferenza, nel presupposto che una regola di inferenza correttamente individuata permetta di potenziare la capacità di elaborare conclusioni da premesse e quindi l'efficacia della condotta umana nel perseguimento di obiettivi.

Ma quali sono le modalità di conoscenza indiretta che le regole di inferenza possono ricostruire e potenziare? Dalla risposta a questa domanda derivano conclusioni significative riguardo l'intelligenza artificiale e le sue prospettive.

Una premessa fondamentale della conoscenza indiretta è la *descrizione del mondo* ovvero, se vogliamo esser più specifici, l'individuazione dello *spazio di problemi* all'interno del quale determinare forme di condotta efficace per il perseguimento di obiettivi. I *criteri di visualizzazione* sono di importanza centrale nella costruzione di descrizioni a partire da un insieme dato di osservazioni. A questo riguardo, il concetto di *frame* (che potremmo considerare una sorta di "quadro associativo") fa riferimento ad un'operazione cognitiva originaria e fondamentale. Il linguista Charles Fillmore definisce questo concetto nei termini di un "gruppo di parole" e osserva: «[C]iò che tiene assieme questi gruppi di parole è il fatto di essere motivati da, fondati su, e strutturati assieme a, specifici quadri unificanti della conoscenza, ovvero schematizzazioni coerenti dell'esperienza, per i quali possiamo usare il termine generico di *frame*» (Fillmore 1985, p. 223; si veda anche Minsky 1975). Il tessuto connettivo alla base del concetto di *frame* è l'*attivazione simultanea* delle parole collegate ad una data sfera dell'esperienza umana, dove il criterio associativo *precede* i singoli termini fra loro connessi all'interno di ciascun *frame*. Qui interviene una proprietà già considerata nel *System of Logic* di John Stuart Mill, secondo il quale «[è] evidente [...] che se consideriamo due nomi fra loro correlati, [come] *padre* e *figlio*, ad esempio, benché gli oggetti denotati dai nomi siano diversi, tuttavia entrambi, in un certo senso, connotano la stessa cosa» (Mill 1846, p. 29; citato in Fillmore 1985, p. 224).

A questo proposito, Fillmore osserva: «L'idea di una presupposta struttura di relazioni (il *fundamentum relationis*), rispetto alla quale parole come *padre* e *figlio* vanno comprese, è senz'altro simile alla nozione di *frame* semantico: noi possiamo conoscere i significati delle parole individuali solo se prima di tutto comprendiamo la base fattuale della relazione che quei termini identificano» (Fillmore 1985, p. 224).

Il concetto di “*frame* semantico” presuppone un contesto pragmatico di carattere associativo. Al tempo stesso il criterio associativo permette di collegare termini che fanno riferimento alla stessa sfera di esperienza, vale a dire termini che si riferiscono alla «storia di eventi che crearono la relazione denominata da ciascuno di quei termini» (Fillmore 1985, p. 224; si veda anche Fillmore 2006). In questo modo si rivela il collegamento tra *frame* semantici e spazi concettuali, dove questi ultimi possono essere considerati spazi di associazioni cognitive generate da *relazioni virtuali* che preesistono ai termini posti in relazione⁽¹⁾. Come osserva Gilles Fauconnier, gli spazi concettuali (da lui denominati “spazi mentali») «non sono parte del linguaggio stesso [...], ma il linguaggio non può esistere senza di essi» (Fauconnier 1985, p. 1).

Nella condotta umana orientata al perseguimento di obiettivi, la capacità di visualizzazione è strettamente collegata alla capacità di selezionare aspetti rilevanti tratti dall'universo di informazioni disponibili. È questo il tema che Herbert Simon colloca al centro della sua analisi delle *descrizioni semplificate efficaci* di un sistema complesso:

Se chiediamo ad una persona di disegnare un oggetto complesso – ad esempio un volto umano – questa persona procederà quasi sempre secondo un criterio gerarchico. Prima di tutto disegnerà i contorni del volto. Poi aggiungerà aspetti come occhi, naso, bocca, orecchi, occhi. Se le chiediamo un maggior grado di dettaglio, inizierà ad inserire particolari per ciascuno di quegli aspetti, come pupille, palpebre, ciglia per gli occhi, e così via sino a raggiungere i limiti della sua competenza anatomica. La sua informazione riguardo all'oggetto è organizzata gerarchicamente nella sua memoria, allo stesso modo di uno schema sintetico di punti (Simon 1962, p. 477).

In generale, secondo Simon,

quanto complessa o semplice sia una struttura dipende in termini fondamentali dal modo in cui noi la descriviamo. La maggior parte delle strutture complesse che troviamo nel mondo sono enormemente ridondanti, e noi

(1) In prospettiva analoga Margaret Boden ha osservato che uno spazio concettuale costituisce «un sistema generativo di qualche tipo (capace di generare una varietà di strutture)» e che qualcosa potrebbe accadere o non accadere «solo rispetto ad un particolare sistema concettuale» (Boden 1992).

possiamo fare uso di questa ridondanza per semplificare la loro descrizione. Ma per fare uso della ridondanza, per realizzare la semplificazione, dobbiamo trovare la rappresentazione appropriata (Simon, 1962, p. 481).

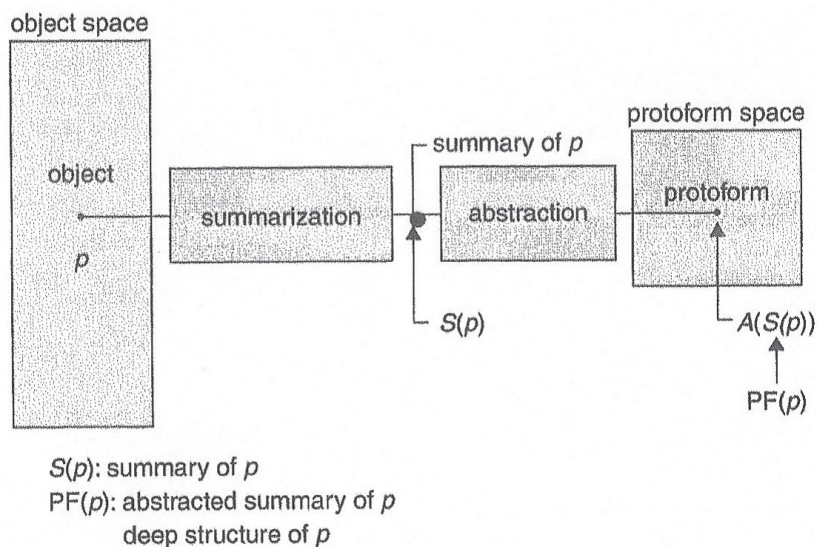
In questa prospettiva, il giudizio sull'efficacia della rappresentazione di una data struttura può essere ricondotto all'orientamento finalistico della condotta umana:

[e]siste un corpus crescente di dati empirici secondo cui l'attività umana [...] di soluzione di problemi è sostanzialmente una forma di analisi mezzi-fini volta a scoprire una descrizione di processo della condotta diretta verso l'obiettivo desiderato (Simon 1962, p. 479).

Un aspetto significativo della selezione di aspetti rilevanti riguarda i *criteri di messa a fuoco* e la *memoria*. Si tratta di questioni collegate fra loro. In entrambi i casi ci troviamo di fronte a processi di *riduzione della complessità* che conducono all'"eliminazione" di aspetti della struttura considerati pragmaticamente irrilevanti rispetto al tipo di condotta in oggetto. Per quanto riguarda la messa a fuoco, la *teoria dei prototipi* elaborata da Eleanor Rosch propone un quadro interpretativo centrato sul ruolo delle categorie e sulla loro utilizzazione nei processi di visualizzazione delle strutture (Rosch 1975, 1981; Rosch e Mervis 1975). Nella teoria dei prototipi, gli attori non elaborano categorie sulla base dell'enumerazione di caratteristiche ma attivano nei contesti pragmatici il prototipo di ciascuna categoria e su quella base assegnano oggetti a questa o a quella categoria a seconda della maggiore o minore distanza di quegli oggetti rispetto ad un particolare prototipo. Questo punto di vista richiama la teoria matematica degli "insiemi sfumati" (*fuzzy sets*) proposta da Lotfi A. Zadeh e in particolare il concetto di 'protoforma' (*protoform*) alla base di quella teoria. Questo concetto è definito da Zadeh in questi termini:

[i]l termine "protoforma" è un'abbreviazione di "forma prototipica". In termini informali, una protoforma, A , di un oggetto, B , scritta come $A = PF(B)$, è un sommario astratto di B [...] Abitualmente, B è una proposizione, un sistema di proposizioni, una domanda, un comando, uno scenario, un problema decisionale, e così via. In termini generali, B può essere una relazione, un sistema, un caso, una forma geometrica, o un oggetto di arbitraria complessità [...] La principale funzione di $PF(B)$ è mettere in evidenza la struttura semantica profonda di B (Zadeh 2011, p. 129).

In termini grafici, Zadeh presenta il concetto di 'protoforma' con questa figura (Zadeh 2011, p. 130):



Il riferimento a una particolare protoforma, ovvero la sua attivazione, nella strutturazione di un insieme dato di informazioni richiama da vicino i processi attraverso i quali si realizzano sia l'attivazione temporale della memoria (basata sulla distinzione fra sistemi a memoria breve e sistemi a memoria lunga; Volterra 1909; Caputo 2019; Quadrio Curzio e Scazzieri 2025)⁽²⁾ sia il ricorso alla dimenticanza come criterio di attivazione di memorie selettive (Assmann 2020; Caputo 2020; Scazzieri 2020)⁽³⁾.

(2) Nella Nota *Sulle equazioni integro-differenziali della teoria dell'elasticità*, presentata ai Lincei nella seduta della Classe di Scienze Fisiche, Matematiche e Naturali del 7 novembre 1909, Volterra scrive: «Come ha osservato il sig. Picard nel suo bell'articolo su: *La mécanique classique et ses approximations successives* [Picard 1907] la Meccanica può distinguersi in meccanica dell'ereditarietà e in meccanica della non ereditarietà. Quest'ultima contempla il caso in cui l'avvenire di un sistema non dipende in un dato istante che dal suo stato attuale o tutt'al più dallo stato infinitamente vicino che precede; la prima riguarda invece quei casi in cui ogni azione lascia una eredità nel sistema e lo stato attuale dipende da tutta la storia precedente [...] Le questioni da lungo tempo studiate di isteresi, di elasticità susseguente, [...] di trainage, rientrano nella meccanica della ereditarietà o per dir meglio nella Fisica della ereditarietà» (Volterra 1909, p. 296).

(3) Aleida Assmann ha osservato che le “propensioni alla dimenticanza” possono attivare oppure disattivare la salienza di modelli culturali e definizioni di identità collettiva (Assmann 2020). Michele Caputo ha sottolineato il ruolo fondamentale delle diverse durate della memoria nella dinamica differenziata dei sistemi economici, e la possibilità di considerare cambiamenti nella durata di memoria come indicatori di cambiamento strutturale in quei sistemi (Caputo 2020).

In questa prospettiva, la capacità di integrare fra loro diversi *frames/spazi* concettuali può essere di importanza decisiva quando si tratta di orientare la condotta in contesti caratterizzati dalla pluralità di attori e punti di vista. A questo proposito è stato osservato che processi di integrazione (*blending*) possono dare luogo alla combinazione di spazi concettuali talora anche in contrasto reciproco ma in grado di generare, appunto attraverso la loro ibridazione, uno spazio concettuale più generale elaborato immaginando prossimità e associazioni estranee agli spazi concettuali di partenza. Mark Turner ha osservato a questo proposito:

Uno spazio mentale integrato ha spazi di immissione (*input spaces*). Esiste una proiezione parziale dagli spazi di immissione allo spazio integrato [...] In modo cruciale, gli spazi integrati possono generare a loro volta una struttura emergente, e possono proiettare questa struttura *all'indietro* verso i propri spazi di immissione. Gli spazi di immissione possono essere non solo *fornitori* di proiezioni allo spazio integrato, ma anche *ricevitori* di proiezioni dallo spazio integrato che essi hanno generato (Turner 1996, p. 60).

Un'interessante proprietà degli "spazi integrati" (*blended spaces*) è il loro collegamento con la visualizzazione di proiezioni, e quindi con immaginazione e creatività:

Benché lo spazio integrato sarà conforme alla propria logica, esso è libero dai vincoli di possibilità che limitano gli spazi generatori [*input spaces*]. Per mezzo di questi aspetti specifici propri di entrambi gli spazi generatori, lo spazio integrato può attivare in modo potente entrambi gli spazi e mantenerli facilmente attivi mentre svolgiamo lavoro cognitivo per costruire significato. La mente può fondarsi su questo insieme di informazioni attive per sviluppare una proiezione (Turner 1996, p. 61).

"ECONOMIE ARTIFICIALI" FRA STRUTTURA, REGOLE DI CONDOTTA E DINAMICHE EVOLUTIVE

L'analisi economica è caratterizzata dalla costruzione di "economie artificiali" sin dalla fase formativa delle sue teorie (Pasinetti 1989 [1986]; Baranzini e Scazzieri 1986). Questo se intendiamo per "economia artificiale" una stilizzazione del sistema economico costruita a fini euristici mettendo in evidenza alcuni aspetti e relazioni al posto di altre. Questo processo si è svolto attraverso due principali modalità.

Nel primo caso, proprio degli economisti fisiocratici e classici, e degli autori successivi ad essi collegati, si parte da una rappresentazione della strut-

tura di interdipendenze fra componenti del sistema economico (attori, settori produttivi, classi sociali) e si studiano le conseguenze di questa struttura sulle modalità di funzionamento del sistema complessivo (Pasinetti 1975, 1980 [1973]). I criteri di condotta umana sono subordinati alla posizione relativa delle componenti all'interno dello schema di interdipendenze e possono, in linea di principio, modificarsi a seconda della rappresentazione strutturale adottata nel definire questa condotta (Cardinale 2018 a, b, c).

Nel secondo caso, caratteristico degli economisti marginalisti e “neo-classici”, si parte dalle regole di condotta, in genere considerate come regole di inferenza o regole “di soglia” (*threshold*), e si va alla ricerca di economie artificiali generate dall'interazione fra attori le cui azioni sono determinate dall'applicazione di quelle regole. La versione standard di questo approccio va alla ricerca di configurazioni di equilibrio del sistema economico (nel senso di configurazioni che rendano possibile la reciproca coerenza delle azioni dei diversi attori), considerando queste configurazioni come individuate da regole di inferenza in grado di determinare le azioni degli attori in modo puntuale date le informazioni disponibili sullo stato del mondo e condizioni di persistenza strutturale (Walras 1874-1877; Arrow 1951; Debreu 1959; Arrow e Hahn 1971). Versioni successive di questo approccio introducono attori (eventualmente eterogenei fra loro) le cui azioni sono guidate da regole di condotta “ad esito aperto” (*open-ended*), e possono dare origine a sentieri di cambiamento strutturale endogeno (morfogenesi) (Simon 1983; Nelson e Winter 1982; Arthur 1994; Dosi 2023). A questa linea di indagine appartengono le “economie artificiali” costruite a partire da paradigma e strumenti dell'intelligenza artificiale applicati ai processi cognitivi degli attori economici. In questo caso, «un agente artificiale è un'entità capace di elaborare le informazioni mediante l'esecuzione di un algoritmo, vale a dire mediante l'applicazione sequenziale di un insieme di regole» (Mercado 2022, p. 10)⁽⁴⁾.

In conclusione, la condotta umana, di cui le azioni economiche sono un caso particolare, genera e utilizza strumenti, come gli apparati tecnici, le regole di inferenza e le regole per la formazione di decisioni. È possibile distinguere fra *strumenti diretti* e *strumenti indiretti*. I primi sono immediatamente rivolti al raggiungimento di obiettivi pratici. I secondi non lo sono sempre, oppure lo sono soltanto in ultima istanza, come accade con la conoscenza scientifica, il cui obiettivo primario è la lettura del “gran libro della natura” (Galileo). Le economie artificiali costruite dall'analisi economica possono essere collegate all'una oppure all'altra di queste due concezioni. Negli econo-

(4) Ruben Mercado richiama l'attenzione sul fatto che «[l']agente artificiale estrae “elementi” di informazione dal proprio ambiente, li elabora al fine di generare una rappresentazione del proprio mondo e, data questa rappresentazione, agisce di conseguenza per realizzare certi obiettivi» (Mercado 2022, p. 10).

misti classici e negli economisti postclassici che a essi si collegano il criterio di lettura del mondo economico attraverso l'analisi di strutture di interdipendenza relativamente persistenti prevale sul criterio di ricostruzione dei criteri di condotta. Quindi le strutture sono in primo piano, mentre le azioni sono prese in considerazione solo in modo indiretto. Invece l'attenzione per le forme di condotta prevale sulla considerazione delle strutture sia dove le strutture sono considerate premessa analitica per il coordinamento fra attori (teorie di tipo walrasiano) sia dove la stessa generazione delle strutture (morfogenesi) è considerata il risultato delle azioni di attori che interagiscono fra loro sulla base di particolari regole di inferenza e di formazione delle decisioni.

COMPLESSITÀ STRUTTURALE FRA VINCOLI DI INTERDIPENDENZA E GRADI DI LIBERTÀ: IL RUOLO DELLA TEORIA ECONOMICA

La relazione fra strutture e complessità è di importanza centrale per chiarire il ruolo della teoria nell'analisi economica (Scazzieri e Quadrio Curzio 2022). Il concetto di *complessità strutturale* fornisce un'utile chiave di lettura a questo riguardo. Si tratta dell'idea che nei sistemi economici la complessità sia organizzata su più livelli, con caratteristiche di *ridondanza* che rendono possibile una pluralità di distinte rappresentazioni della loro struttura. Nelle parole di Herbert Simon: «se una struttura complessa è del tutto non ridondante – se nessun aspetto della sua struttura può essere derivato da qualche altro aspetto – allora quella è la sua descrizione più semplice. La possiamo mostrare ma non la possiamo descrivere facendo riferimento ad una struttura più semplice» (Simon 1962, p. 478). D'altra parte, una configurazione a più livelli della complessità introduce forme di ridondanza che rendono possibili descrizioni semplificate, talora significativamente diverse fra loro. Un esempio illuminante è la possibilità di descrivere le interdipendenze fra attività produttive in un sistema complesso di divisione del lavoro facendo riferimento, in alternativa, a una rappresentazione 'orizzontale' della struttura in termini di flusso circolare (Leontief 1991 [1928]; Sraffa 1960; Pasinetti 1975; Quadrio Curzio 1996) oppure ad una rappresentazione 'verticale' della struttura in termini di una collezione di sub-sistemi verticalmente integrati costruiti a partire dai beni finali di consumo e di investimento (Pasinetti 1980 [1973]) oppure a partire da particolari risorse non prodotte (Quadrio Curzio 1986). Ridondanza e pluralità di rappresentazioni strutturali richiamano l'attenzione sui criteri di visualizzazione delle interdipendenze (Cardinale 2018c; Cardinale e Scazzieri 2023; Pabst e Scazzieri 2023; Scazzieri 2023) e sul ruolo degli spazi concettuali come generatori di *rappresentazioni virtuali* che permettano all'immaginazione teorica di andare oltre distinzioni consolidate per mettere

a fuoco le forme di interdipendenza che via via si modificano lungo le traiettorie di cambiamento strutturale.

Luigi Pasinetti, grande economista Linceo da poco scomparso, ha richiamato l'attenzione sul fatto che:

in economia (e nelle scienze sociali in genere) [esiste] una molteplicità di teorie e di approcci molto più variegata che nelle scienze naturali. I grandi paradigmi della teoria economica, più che rimpiazzarsi, sembrano coesistere nel tempo, anche se con diversa intensità, con vicende alterne. Proprio perché la realtà di un sistema economico è in continuo movimento, non sorprende che più teorie cerchino di coglierne e interpretarne, attraverso la ricerca, aspetti diversi, certe volte addirittura contrastanti, anche se tutti in qualche modo rilevanti (Pasinetti 2013, pp. 47-48).

Un importante aspetto di queste osservazioni è l'enfasi sull'efficacia euristica del lavoro di *costruzione di teorie*⁽⁵⁾. Possiamo aggiungere, anche alla luce di quanto si è detto in precedenza, che lo stesso lavoro teorico trova, in ultima analisi, un significativo fondamento proprio nell'apertura alla *molteplicità di teorie*, dal momento che questa molteplicità esprime concretamente la *pluralità di possibili rappresentazioni* di una struttura di interdipendenze e l'apertura della teoria ai cambiamenti strutturali, a volte epocali, che hanno luogo nella storia. In questa prospettiva, il ruolo della teoria, e più in generale della disposizione verso la scienza pura, nell'analisi economica è soprattutto quello di orientare l'immaginazione all'interno di spazi concettuali per descrivere e spiegare le complesse interdipendenze che caratterizzano la vita economica e le loro trasformazioni nel tempo.

(5) In prospettiva diversa, ma vicina a questo punto di vista, Vito Volterra insiste sulla necessità che le costruzioni teoriche procedano senza ostacoli lungo percorsi che possano condurre alla scoperta di *relazioni possibili* non ancora sperimentate in contesti pragmatici. Un'indicazione significativa dell'atteggiamento di Volterra è data dalle sue osservazioni sulle geometrie non Euclidee: «[u]no scritto di carattere schiettamente geometrico del Beltrami, il quale attingeva le sue origini alle ricerche di Gauss, di Lobatschewski e di Riemann sulla geometria non euclidea, fu di tanto universale importanza da rischiare di novella luce la teoria della conoscenza e i fondamenti della logica stessa» (Volterra 1907, p. 232).

BIBLIOGRAFIA

- ARROW 1951 = K.J. ARROW, *An Extension of the Basic Theorems of Classical Welfare Economics*, in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, a cura di J. Neyman, Berkeley, 1951, pp. 507-532.
- ARROW, HAHN 1971 = K.J. ARROW, F. HAHN, *General Competitive Analysis*, San Francisco and Edinburgh 1971.
- ARTHUR 1994 = W.B. ARTHUR, *Increasing Returns and Path Dependence in the Economy*, Ann Arbor 1994.
- ASSMANN 2020 = A. ASSMANN, *Global Publics – Remembering and Forgetting*, «Balzan Papers» III, 2020, pp. 127-146.
- BARANZINI, SCAZZIERI 1986 = M. BARANZINI, R. SCAZZIERI, *Knowledge in Economics: A Framework*, in M. Baranzini, R. Scazzieri (a cura di), *Foundations of Economics. Structures of Inquiry and Economic Theory*, Oxford and New York, 1986, pp. 1-87.
- BODEN 1992 = M. BODEN, *Creativity and Computers*, Lecture delivered at the workshop *Brains as Machines and Machines as Minds*, Cambridge, Cambridge Philosophical Society, 7 December 1992.
- BOLZONI 2003 = L. BOLZONI, *Le tecniche della memoria e la costruzione degli spazi interiori fra Medioevo e Rinascimento*, «Lettere italiane» 55 (2003), 1 (gennaio-marzo), pp. 26-46.
- CAPUTO 2019 = M. CAPUTO, *Memory, Mathematics and Economics*, Accademia Nazionale dei Lincei, *Memorie*, Classe di Scienze Morali, Storiche e Filologiche, serie 9, volume 40 (2019), fascicolo 1.
- CAPUTO 2020 = M. CAPUTO, *Memory, Mathematics, and Economy*, «Balzan Papers» III (2020), pp. 147-154.
- CARDINALE 2018a = I. CARDINALE, *Structural Political Economy*, in I. Cardinale, R. Scazzieri (a cura di), *The Palgrave Handbook of Political Economy*, London, 2018, pp. 769-784.
- CARDINALE 2018b = I. CARDINALE, *Classical Political Economy*, in I. Cardinale, R. Scazzieri (a cura di), *The Palgrave Handbook of Political Economy*, London, 2018, pp. 171-189.
- CARDINALE 2018c = I. CARDINALE, *A Bridge over Troubled Water: A Structural Political Economy of Vertical Integration*, «Structural Change and Economic Dynamics» 46 (September 2018), pp. 172-179.
- CARDINALE, SCAZZIERI 2023 = I. CARDINALE, R. SCAZZIERI, *First- and Third-Person Viewpoints in Economic Theory*, Cambridge, 2023.
- DEBREU 1959 = G. DEBREU, *Theory of Value: an Axiomatic Analysis of Economic Equilibrium*, New Haven and London 1959.
- DOSI 2023 = G. DOSI, *The Foundations of Complex Evolving Economies: Part One: Innovation, Organization, and Industrial Dynamics*, Oxford 2023.
- FAUCONNIER 1985 = G. FAUCONNIER, *Mental Spaces*, Cambridge, Mass. 1985.

- FILLMORE 1985 = C. FILLMORE *Frames and the Semantics of Understanding*, "Quaderni di semantica" 6, 2 (dicembre 1985), pp. 222-254.
- FILLMORE 2006 = C. FILLMORE, *Invoking and Evoking Frames*, relazione presentata al convegno *Frames. A Colloquium in Linguistics, Philosophy and Economics*, Alma Mater Studiorum Università di Bologna, 6 giugno 2006.
- LEONTIEF 1991 [1928] = W. LEONTIEF, *The Economy as a Circular Flow*, «Structural Change and Economic Dynamics» 2, 1 (June 1991), pp. 181-212 (pubblicato in origine come *Die Wirtschaft als Kreislauf*, «Archiv für Sozialwissenschaft und Sozialpolitik» 60 (1928), pp. 577-623).
- MERCADO 2022 = R. MERCADO, *Artificial Economics. Methods, Models, and Interdisciplinary Links*, Cambridge 2022.
- MILL 1846 = J. S. MILL, *A System of Logic, Ratiocinative and Inductive: being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*, seconda edizione, London 1846.
- MINSKY 1975 = M. MINSKY, *A Framework for Representing Knowledge*, in P.H. Winston (a cura di), *The Psychology of Computer Vision*, New York, 1975, pp. 211-277.
- NELSON, WINTER 1982 = R.R. NELSON, S.G. WINTER, *An Evolutionary Theory of Economic Change*, Cambridge (Massachusetts) and London 1982.
- PABST, SCAZZIERI 2023 = A. PABST, R. SCAZZIERI, *The Constitution of Political Economy. Polity, Society and the Common Weal*, Cambridge 2023.
- PASINETTI 1975 = L.L. PASINETTI, *Lezioni di teoria della produzione*, Bologna 1975.
- PASINETTI 1980 [1973] = L.L. PASINETTI, *The Notion of Vertical Integration in Economic Analysis*, in L.L. Pasinetti (a cura di), *Essays on the Theory of Joint Production*, London and Basingstoke, 1980, pp. 16-43 (pubblicato in origine in «Metroeconomica» 25 (1973), pp. 1-29).
- PASINETTI 1989 [1986] = L. L. PASINETTI, *La teoria del valore come fonte di paradigmi alternativi nell'analisi economica*, in L. L. Pasinetti (a cura di), *Aspetti controversi della teoria del valore*, Bologna, 1989, pp. 231-254. Edizione originale in M. Baranzini, R. Scazzieri (a cura di), *Foundations of Economics. Structures of Inquiry and Economic Theory*, Oxford e New York, 1986, pp. 409-431.
- PASINETTI 2013 = L. L. PASINETTI, *Prospettive e limiti dell'economia quantitativa*, Istituto Lombardo - Accademia di Scienze e Lettere. Incontri di Studio, 2013, pp. 43-55.
- PICARD 1907 = E. PICARD, *La mécanique classique et ses approximations successives*, «Rivista di Scienza» 1, 1 (1907), pp. 4-15.
- QUADRIO CURZIO 1986 = A. QUADRIO CURZIO, *Technological Scarcity: An Essay on Production and Structural Change*, in M. Baranzini, R. Scazzieri (a cura di), *Foundations of Economics. Structures of Inquiry and Economic Theory*, Oxford and New York 1986, pp. 311-338.
- QUADRIO CURZIO 1996 = A. QUADRIO CURZIO, *Production and Efficiency with Global Technologies*, in M. Landesmann, R. Scazzieri (a cura di), *Production and Economic Dynamics*, Cambridge, 1996, pp. 105-126.

- QUADRIO CURZIO, SCAZZIERI 2025 = A. QUADRIO CURZIO, R. SCAZZIERI, *Volterra Presidente dei Lincei: interdisciplinarietà di un genio*, in *Vito Volterra nel centenario della sua Presidenza Lincea*, Accademia Nazionale dei Lincei, Atti dei Convegni Lincei, vol. 360, pp. 63-75.
- ROSCH 1975 = E. ROSCH, *Cognitive Reference Points*, «Cognitive Psychology» 7 (1975), pp. 532-547.
- ROSCH 1981 = E. ROSCH, *Prototype Classification and Logical Classification: The Two Systems*, in E. Scholnick (a cura di), *New Trends in Conceptual Representation*, Hillsdale, N.J., 1981 pp. 73-85.
- ROSCH, MERVIS 1975 = E. ROSCH, C.B. MERVIS, *Family Resemblance: Studies in the Internal Structure of Categories*, «Cognitive Psychology» 7 (1975), pp. 573-605.
- ROSSI 1960 = P. ROSSI, *Clavis universalis: arti mnemoniche e logica combinatoria da Lullo a Leibniz*, Milano e Napoli, 1960; seconda edizione, Bologna, 1983. Edizione inglese: *Logic and the Art of Memory: The Quest for a Universal Language*, London, 2006.
- SCAZZIERI 2020 = R. SCAZZIERI, *Remembering, Forgetting, and the Construction of Identity*, «Balzan Papers» III (2020), pp. 155-160.
- SCAZZIERI 2023 = R. SCAZZIERI, *Sull'interdipendenza strutturale*, Accademia Nazionale dei Lincei, *Lectio brevis*, 12 maggio 2023.
- SCAZZIERI, QUADRIO CURZIO 2022 = R. SCAZZIERI, A. QUADRIO CURZIO, *Introduzione. Sulla complessità teorica e storica dell'economia politica*, in A. Quadrio Curzio, *Economia, complessità, sviluppo. Scritti vari*, Atti della Accademia Nazionale dei Lincei, Classe di Scienze Morali, Storiche e Filologiche, *Memorie*, serie IX, volume XLV, fascicolo 1 (2022), pp. 13-29.
- SIMON 1962 = H. A. SIMON, *The Architecture of Complexity*, in «Proceedings of the American Philosophical Society» 106 (1962), pp. 467-482.
- SIMON 1983 = H.A. SIMON, *Reason in Human Affairs*, Stanford, CA. 1983.
- SRAFFA 1960 = P. SRAFFA, *Produzione di merci a mezzo di merci. Premesse a una critica della teoria economica*, Torino 1960.
- TURNER 1996 = M. TURNER, *The Literary Mind*, New York e Oxford 1996.
- VOLTERRA 1907 = V. VOLTERRA (1907) *Il momento scientifico presente e la nuova Società italiana per il progresso delle scienze*, «Rivista di scienza. Organo internazionale di sintesi scientifica» II (1907), pp. 225-237.
- VOLTERRA 1909 = V. VOLTERRA, *Sulle equazioni integro-differenziali*, «Rendiconti» delle sedute, Reale Accademia dei Lincei, Classe di Scienze Fisiche, Matematiche e Naturali, serie quinta, vol. XVIII (1909), I semestre, pp. 167-174.
- WALRAS 1874-1877 = L. WALRAS, *Économie politique pure*, Lausanne 1874-1877.
- YATES 1966 = F. YATES, *The Art of Memory*, London 1966.
- ZADEH 2011 = L.A. ZADEH, *Generalized Theory of Uncertainty: Principal Concepts and Ideas*, in S. Marzetti Brandolini, R. Scazzieri (a cura di), *Fundamental Uncertainty*, London 2011, pp. 104-150.