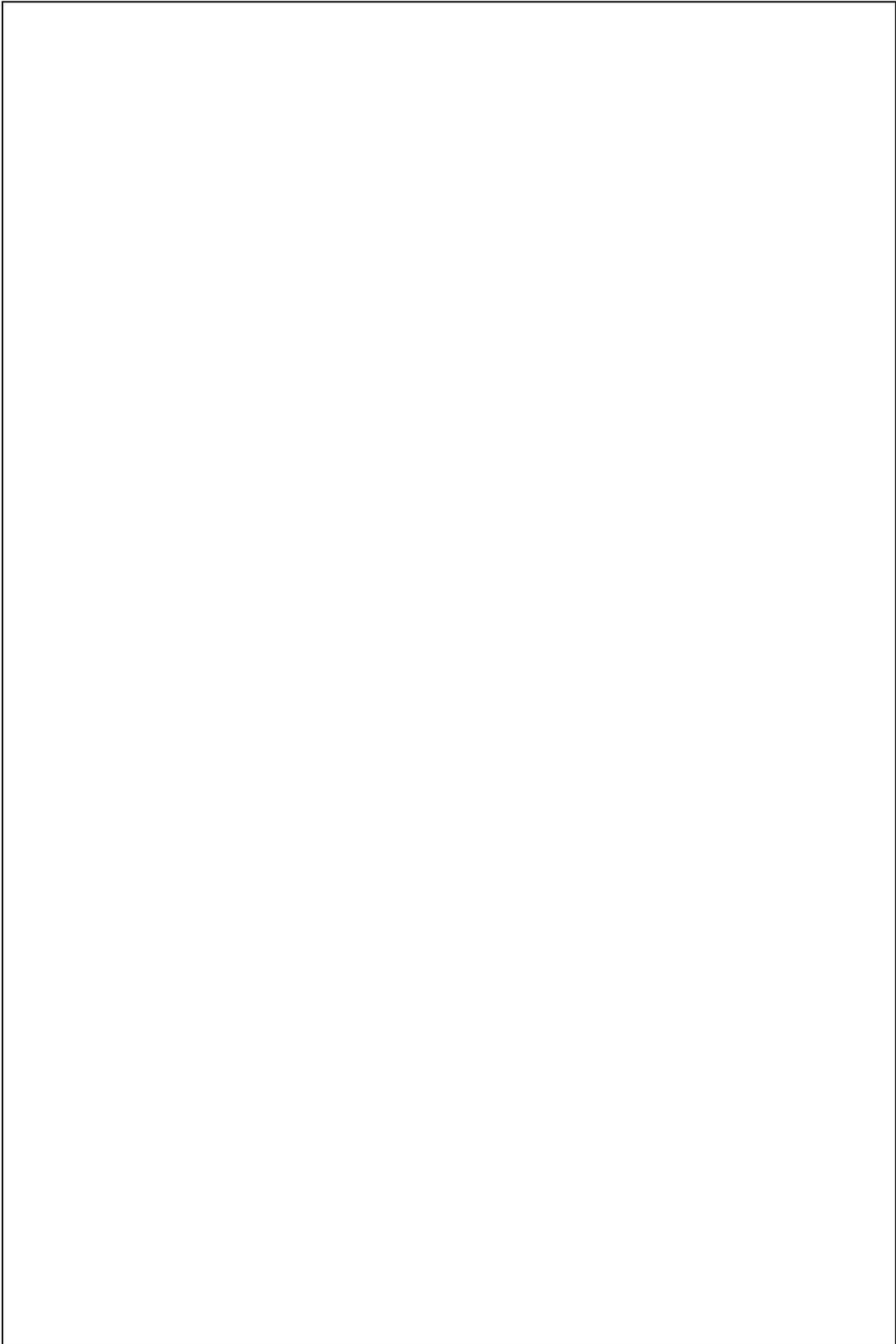


unicollege
working
papers
unicollege
working
papers
unicollege
working
papers
unicollege
working
papers



**Centro
Editoriale
Accademico**
unicollege

unicollege
Knowledge
and Experience.

 **International
Council**[®]

 **Accademia
di Italiano**[®]

 **Adiuva**[®]

**Centro
Editoriale
Accademico**
unicollege

Via Bolognese 52

50139 Firenze

<https://www.unicollegessml.it/centro-editoriale-accademico/>
centro.editoriale.accademico@unicollegessml.it

Unicollege Working Papers

Collana diretta da Lorenzo Grifone Baglioni

- 1 *Sociogenesi dell'Intelligenza Artificiale*, Andrea D'Angelo.
- 2 *La persona al centro*, Elisa Gallocchio, Barbara Bonomi.

How to cite this paper / Come citare questo saggio:

Andrea D'Angelo (2025), *Sociogenesi dell'Intelligenza Artificiale*, "Unicollege Working Papers", 1, 1, 7-42.

Unicollege Working Papers

Volume 1, Issue 1

Centro Editoriale Accademico - Firenze

ANNO 2025 - ISSN

Andrea D'Angelo

Sociogenesi dell'Intelligenza Artificiale

Abstract: Artificial Intelligence (AI) is no longer just the project of a powerful tool, but a generative reality with significant social implications. Since its inception, AI has been characterized by an experimental and branching evolution, driven in at least two directions: horizontally, in application fields often only partially known to the public, and vertically, with a surprising increase in computational potential and bush-like generative speciation. The essay addresses some milestones in AI's historical development, from foundational principles to early research programs, from Expert Systems to the development of neural networks and LLMs, proposing a socio-genesis of this new and promising social actor.

Keywords: Artificial Intelligence, Evolution, Historical Development, Expert Systems, Neural Networks.

Contributor: Graduated in Theoretical Philosophy (BA) and Logic, Philosophy and History of Science (MA), Subject Expert in Sociology of Change, Course Teacher of Philosophy of Mind and Artificial Intelligence and its implementation in education at Unicollege SSML <andrea.dangelo@adiuva.it>.

1. Introduzione

Poche cose viaggiano veloci come le rivoluzioni, reali o supposte, sociali o scientifiche. Non è raro poi che non ci si renda conto di trovarsi sul limitare di un punto di svolta e si proceda spediti sotto la spinta di un entusiasmo genuino, a tratti

infantile. Sarà solo un successivo senso di scomodità e spaesamento a sollevare i fondati dubbi che il mondo nel quale si sta vivendo non è - e non può più essere, quello a cui si era abituati.

Quando a partire dagli anni Ottanta il matematico e scrittore Vernor Vinge introduce il termine “singolarità tecnologica”, ad indicare il raggiungimento di un’accelerazione dello sviluppo tecnologico tale da non poter permettere più alla generazione precedente di comprenderne la natura e prevederne le conseguenze, le letture catastrofiste si moltiplicano, la fantascienza ne fiorisce. Sono i miraggi di intelligenze artificiali che in un virtuoso circolo di autosviluppo potenziano le proprie capacità oltre il controllabile e l’umano. Tuttavia, il termine “singolarità” non è da intendersi matematicamente come banalmente foriero di una crescita all’infinito, ma viene preso a prestito dalla fisica come “a point where our models must be discarded and a new reality rules” (Vinge 1993). Tutti i modelli di comprensione e predizione dell’ambiente circostante diventerebbero obsoleti e l’Uomo si troverebbe scaraventato in una nuova fase dello sviluppo della sua civiltà. Benché si sia ancora lontani dal raggiungere quel

momento¹, l'utilizzo dell'Intelligenza Artificiale ha già mutato ambienti, costumi e società in modi difficilmente quantificabili e dei quali il grande pubblico spesso non sospetta neanche l'esistenza. E promette di fare di più.

Le IA sono già in grado di vedere (Parkhi 2015), ascoltare (Hoy 2018), riconoscere le e-mail indesiderate (Goodman *et al.* 2007) o le frodi bancarie (Phua 2010), guidare veicoli (Grigorescu *et al.* 2020), profilare utenti (Smith, Linden 2017), fare diagnosi mediche (De Fauw *et al.* 2018). Solo adesso che l'accelerazione della loro diffusione ha raggiunto velocità superiori alla comune capacità di adattarsi ai cambiamenti, spostandosi in campi nei quali non si prevedeva l'invasione dell'Artificiale, l'urgenza di conoscere questo nuovo soggetto si è fatta palpabile, i timori più concreti.

Anche lontani dalla singolarità, si ha l'impressione che lo sviluppo dell'Intelligenza Artificiale abbia tutte le caratteristiche per affrancarsi dal semplicistico concetto di strumento. Ciò che promette di raggiungere è una quantomeno parziale autonomia funzionale nello sviluppo di compiti di grande importanza per le civiltà moderne e

¹ Vinge (1993) si è notoriamente detto convinto del raggiungimento della singolarità entro il 2030.

soprattutto di farlo permettendo ai suoi stessi sviluppatori di avere solo minima conoscenza del suo esatto stato funzionale in un dato momento. Costruire la “black box” del comportamentismo artificiale non può che segnare il passaggio dallo status di strumento scomponibile e analizzabile a quello di automa-processore di input e output. Il mondo interno delle funzioni dell’Intelligenza Artificiale si sottrarrebbe al controllo e all’analisi, costringendoci a un’interazione per tentativi ed errori, nella speranza di spingere il mondo funzionale della macchina a muoversi verso output comportamentali desiderati.

Ecco allora che l’analisi ingegneristica lascia il posto a quella “etologica”: costruire un’Intelligenza Artificiale siffatta non sarebbe più unicamente progettazione di strumenti, avvicinandosi piuttosto a quello di “allevamento” di entità con le quali per inevitabile convenienza delle società sarà necessario avviare un processo di convivenza – e coevoluzione. Più che una nuova scoperta dell’elettricità, l’Intelligenza Artificiale potrebbe costituire il raggiungimento di una nuova forma di allevamento, non meno rivoluzionaria della prima. La creazione di una mente oscura apre già spiragli di futuro dove i nostri modelli vacillano. In quest’ottica, il presente lavoro vuole essere una

succinta descrizione dello sviluppo filogenetico dell'animale artificiale del quale stiamo con tanto entusiasmo osservando l'ontogenesi embrionale.

2. Idee e progetti: La fecondazione

Erroneamente si penserebbe all'Intelligenza Artificiale come a una sorta di spontaneo completamento teleologico delle scienze informatiche. Lo sviluppo delle IA non segue la trama classica di un'epica preconfezionata: non c'è una biografia che accompagni il lettore dai primi passi fino al massimo sviluppo, descrivendo avvio, crescita e arrivo di qualcosa già destinato a cambiare tutto per sempre. Non lo è nella forma, susseguirsi di ciclici declini e rinascite; non lo è nei contenuti, fin da subito ben consci di quali vie valesse la pena battere. Persino le Reti Neurali, ormai la base degli algoritmi IA più moderni, non costituirono un lampo a ciel sereno, insperato e inaspettato; al contrario, fecero la loro comparsa pressoché negli stessi anni in cui venivano avanzati i primi progetti di computer digitali.

Lungi dall'essere una carta tardiva tirata fuori dal mazzo contro ogni previsione e aspettativa, furono invece il frutto di un lento processo di ricerche e fallimenti precocemente avviato, spesso

segnato dalla delusione. Questo perché il collegamento tra macchine intelligenti e meccanicizzazione del pensiero umano (e animale, più in generale) costituisce un unico fenomeno dalla doppia faccia. Nella costruzione di una macchina che pensasse crescevano dopotutto anche le speranze mai sopite della comprensione del pensiero razionale stesso. Scienze neurologiche e scienze informatiche in questo costituirono una sorta di sistema binario in autosostentamento.

Quando nel 1936 il matematico inglese Alan Turing pubblica la prima formulazione (Turing 1936) delle sue famose Macchine computazionali, antesignane dei moderni computer, quella che andava a posizionare non era solo la pietra d'angolo delle scienze informatiche e della teoria della computabilità. Il suo lavoro si inseriva all'interno del fervore che ruotava attorno all'ambizioso programma di Hilbert, che mirava alla dimostrazione rigorosa della consistenza dei sistemi formali. In particolare, l'Entscheidungsproblem (il problema della decisione) costituiva forse il punto cardine di tutta la ricerca matematica e si riassumeva nell'annosa questione del dimostrare l'esistenza di un metodo generale per determinare se un'affermazione fosse deducibile utilizzando i soli stru-

menti della logica. Risolvere positivamente il problema significava trovare un metodo universale e replicabile per la risoluzione di qualunque quesito matematico, il Santo Graal della logica formale. La cosa richiedeva per la prima volta nella storia di concettualizzare chiaramente la definizione di quesito matematico sensato, ovvero effettivamente calcolabile e dotato di una risposta definita, determinata e conclusiva. Fuori del tracciato del programma di Hilbert, destinato all'oblio², quello che Turing lasciava al mondo con il suo articolo era non solo il rudimentale modello teorico di un computer, ma la definizione stessa di "quesito matematico sensato" ovvero di "effettivamente calcolabile" come risolvibile da una delle macchine sulle quali di lì a poco avrebbe iniziato a lavorare concretamente.

Sebbene non fosse stato il primo ad avanzare una definizione chiara di "effettivamente calcolabile", così come non fu il primo a dimostrare l'impossibilità di risolvere l'Entscheidungsproblem³, vi è tuttavia un prima e un dopo Turing. Il

² Tanto dal punto di vista della completezza e coerenza, grazie ai lavori di Godel, quanto dal punto di vista della decidibilità, grazie alle dimostrazioni contemporanee di Church e Turing sull'insolubilità del problema di Hilbert.

³ Lo anticipò di qualche mese Alonzo Church (1936) utilizzando il suo λ -calcolo di recente invenzione per dare una definizione

progetto di una macchina in grado di assolvere al ruolo di computer perfetto fugava ogni pretesa di monopolio della mente umana sul mondo della conoscenza razionale. La dimostrazione dell'esistenza di una macchina in grado di arrivare per definizione fino ai limiti umani – di più, ai limiti della matematica stessa – tirava la mente umana giù dall'olimpico dell'irraggiungibile e apriva la strada a una deduzione forzata e seducente: se c'era modo di ridurre un qualsiasi processo a un algoritmo, per quanto complesso fosse, allora in linea di principio quel processo era simulabile da una macchina – meccanicisticamente comprensibile.

Di lì a poco l'intera ricerca globale avrebbe subito una trasformazione di paradigma avviando una prolifica, variegata e convinta esplorazione di metodi per ridurre i processi più svariati ad algoritmi: lunghe e intricate sequenze di calcoli. Come fuoco di Prometeo, gli uomini avevano trovato un modo per domare un mondo fino a quel momento inaccessibile. Non dovrebbe sorprendere allora la precocità che si ebbe nel muovere questa nuova, raffinata arma di riduzionismo verso la stessa mente umana.

rigorosa di funzione “effettivamente computabile”, nonché dimostrare l'impossibilità di un metodo che garantisca a priori la decidibilità di un'affermazione matematica qualunque.

Non erano passati dieci anni dalla formulazione delle Macchine di Turing che i principi ad essi sottesi furono mossi per spiegare il funzionamento del cervello umano (McCulloch, Pitts 1943). L'idea era in realtà semplice e riposava tanto sulle nuove idee sollevate da Turing quanto sulle ultime scoperte delle neuroscienze sul funzionamento del cervello. Se, come sembrava, il comportamento dei singoli neuroni cerebrali poteva essere assimilato a quello di un interruttore biologico costretto nei limiti della logica binaria a dividersi tra uno stato elettricamente inerte e uno stato eccitato, allora un algoritmo poteva governarne il comportamento. Fin dalle prime righe del loro studio, le potenzialità di questo approccio erano chiaramente tracciate: "Because of the 'all-or-none' character of nervous activity, neural events and the relations among them can be treated by means of propositional logic" (McCulloch, Pitts 1943, 115). Significava riuscire a schiacciare la complessità del funzionamento del cervello umano dentro i limiti di una semplice computazione ad output binari.

Non è un caso che, quando venne presentato l'architettura del primo calcolatore digitale della storia, le varie componenti logiche della macchina erano dette "correspond to the associative neurons in the human nervous system" (Von Neumann

1945, 3). Se i neuroni avevano due stati, emettere o non emettere una scarica elettrica, dipendenti dagli input ricevuti, allora tutta la loro natura era comprimibile nei termini di una trasformazione di quegli input in uno dei due stati: un algoritmo. E viceversa. Trovato un modello esplicativo, se ne dedusse immediatamente un progetto simulativo: la storia delle reti neurali aveva fatto il primo passo.

Era il 1955 quando John McCarthy, giovane dottore in matematica, entrava come professore assistente al Dartmouth College. Benché ventottenne, aveva già avuto modo di lavorare come istruttore a Princeton e conoscere autori come Claude Shannon, padre della teoria matematica della comunicazione, col quale aveva inoltre partecipato alla stesura di una raccolta di lavori sulla teoria degli automi che tradiva la sua grande passione per le macchine intelligenti. Non sorprende quindi che, giunto al Dartmouth College, si volse spontaneamente alla preparazione di un incontro di ampio respiro sul tema delle macchine intelligenti. Psicologi, matematici ed ingegneri, i partecipanti dell'evento erano mossi più dalla speranza comune per la simulazione della cognizione umana che da una convergenza di metodi di ricerca. I temi proposti furono pressoché i medesimi

che ci si aspetterebbe di trovare anche oggi in un evento simile: i limiti tecnici dei calcolatori artificiali, la possibilità per un computer di simulare il linguaggio umano, reti neurali, sistemi di automiglioramento delle macchine, capacità di astrazione e manifestazione di creatività. Lo stesso termine “Intelligenza Artificiale” fu coniato in questo momento, nel tentativo di fondare il campo di ricerca come autonomo rispetto a teorie analoghe quale ad esempio la cibernetica⁴ (McCarthy 1955).

Il workshop fu vivace, poco coerente, a tratti deludente, e attraversato da problemi organizzativi, di reperimento di fondi e di coesistenza delle varie strategie proposte, cionondimeno costituì il momento di nascita ufficiale del campo dell’Intelligenza Artificiale come di quella “scienza e ingegneria di creare macchine intelligenti” (McCarthy 2004, 2). Il tema fu approcciato e proposto in modo

⁴ La cibernetica, campo multidisciplinare avviato da Norbert Wiener negli anni Quaranta, studiava i sistemi biologici sotto l’analisi della loro autoregolazione in seguito a feedback sui loro comportamenti. L’estensione del concetto aveva portato la cibernetica ad interessarsi del funzionamento delle macchine interagenti con l’ambiente e quindi anche delle macchine intelligenti. Fu anche per evitare un confronto con una personalità forte come quella di Wiener e rischiare che il campo di ricerca ruotasse eccessivamente attorno alla direzione da lui intrapresa che McCarthy conì il termine Intelligenza Artificiale come termine ombrello volutamente ambiguo e aperto alla sperimentazione (Nilsson 2009, 78).

fortemente variegato e sperimentale, esemplificativo di un'impresa di esplorazione di un mondo del quale nessuno aveva davvero un'idea definita, e non vi era – come d'altronde non vi è tutt'ora – una definizione chiara e condivisa di cosa “Intelligenza Artificiale” dovesse davvero significare. Tuttavia, questa mancanza di un'idea chiara e unitaria costituì con tutta probabilità uno dei maggiori punti di forza dell'esplorazione scientifica dell'epoca perché, invece di incanalare le energie verso un'unica strada, contribuì a generare un delta di approcci dalle velleità totalizzanti, ma che più avanti si scoprirono fortemente complementari. La direzione data all'Intelligenza Artificiale a Dartmouth aveva il carattere del fervore e dell'originalità: il concetto di creatività diventava parte integrante del tipo di macchina che si intendeva realizzare.

Il prodotto più notevole dell'esperienza al Dartmouth College fu proprio l'entusiasmo: McCarthy, Marvin Minsky, Allen Newell ed Herbert Simon, quattro grandi pionieri dell'IA, si conobbero, discussero sul tema e avviarono una feconda diaspora di prospettive e centri di ricerca. McCarthy avrebbe fondato lo Stanford Artificial Intelligence Project, Minsky avrebbe avviato il Laboratorio di Intelligenza Artificiale del MIT e Allen Newell

ed Herbert Simon avrebbero lavorato ai primi programmi di intelligenza artificiale mai scritti con tanto successo da spingere Simon a predire che “machines will be capable, within twenty years, of doing any work that a man can do” (Simon 1965, 96).

3. Approcci AI: La segmentazione

Fin dal principio erano identificabili due ben differenti famiglie di strategie per approcciare la disciplina, che potremmo chiamare approcci simbolici e subsimbolici. Gli approcci simbolici, storicamente antecedenti, affondavano le loro radici in un orientamento simulativo delle attività coscienti e più cognitivamente complesse della mente: l'idea era quella di spingere le macchine oltre la simulazione di potenti calcolatrici, per arrivare a fornire loro regole di composizione non banali in modo da renderle in grado di operare in maniera autonoma e architettonicamente complessa su tutto un albero di operazioni sequenziali e oggetti di varia natura logica.

Il primo esempio di questo, precedente persino alla conferenza di Dartmouth, era stato costruito proprio da Newell e Simon alla Rand Corporation nel 1956. Alla macchina fu aperta la possi-

bilità di riconoscere enunciati di logica proposizionale, nonché garantita la possibilità di operare su quegli enunciati attraverso regole di manipolazione degli atomi logici conformi alle leggi fondamentali della logica classica. Nei fatti, il Logic Theorist, questo il nome del primo programma di intelligenza artificiale della storia, era ben più di una calcolatrice di logica proposizionale: la macchina funzionava prendendo ad input degli assiomi forniti dal programmatore e una conclusione logica e cercando tutte le derivazioni possibili che collegassero gli assiomi alle loro deduzioni.

Per fare questo i programmatori dovettero affrontare il problema dell'inevitabile esplosione combinatoria di una macchina non in grado di capire quali derivazioni logiche dagli assiomi avrebbero generato alberi di concatenazioni superflue, se non addirittura circolari. Si trattava di insegnare a una macchina un modo intelligente e "umano" di scartare le sequenze di calcoli che con tutta probabilità avrebbero allungato e complicato inutilmente il procedimento.

Alla macchina furono insegnate strategie di ricerca intelligenti dette "euristiche" come "provare a risalire all'indietro dalla conclusione alle premesse" o "scomporre il problema in parti più semplici e manipolare queste ultime". Di più, la

macchina era voracemente bisognosa di sempre più memoria fisica dove conservare le informazioni ricavate e quelle in uso dal programma: per ovviare al problema furono inventati modi di permettere alla stessa di selezionare informazioni importanti, cancellando dai registri tutte le altre e pulendo la sua memoria fisica, in un processo vagamente ispirato alla memoria associativa umana.

Palesi e preziosi risultano i contributi delle ricerche pregresse di Herbert Simon, premio Nobel per l'economia, sull'analisi del comportamento umano nella scelta di strategie razionali. L'obiettivo principale del gruppo di ricerca non era solo costruire una macchina capace di elaborare enunciati di logica del primo ordine, ma soprattutto riuscire a insegnare a un programma automatico le caratteristiche distintive del ragionamento razionale umano. Il Logic Theorist fu in grado di dimostrare ben trentotto dei principali teoremi logici, in un caso con una dimostrazione formale più elegante di quella degli autori originali. L'Intelligenza Artificiale si era con forza ricavata un suo spazio tra le scienze e, sebbene pochi se ne fossero accorti, tra quelle più promettenti.

Gli approcci simbolici avrebbero promesso e regalato risultati insperati, mostrando le poten-

zialità delle macchine, se computativamente potenti e ben programmate. È degno di interesse come, all'interno del Logic Theorist, del GPS (General Problem Solver, sua evoluzione generalizzata) o in generale dei programmi di ispirazione simbolica, la natura semantica degli oggetti manipolati fosse deliberatamente ignorata: ciò che importava ai fini della procedura era che gli oggetti simbolici fossero manipolabili attraverso chiare regole combinatorie e sintattiche.

La macchina non era infatti in grado di distinguere tra combinazioni di oggetti logici come "A Produce B" o "cane gatto gallo" purché le regole di applicazione tra i soggetti A-cane B-gallo e Produce-gatto rimanessero invariate. Quelle quindi che erano le soluzioni individuate da Logic Theorist o da GPS ai vari problemi logici cui erano applicati non erano altro che sequenze di simboli ben formati secondo le regole combinatorie, dove per "simbolo" si intende un'entità logica di interpretazione variabile. Interpretazione che ovviamente doveva essere fornita dall'utente umano. Il passo successivo dei ricercatori interessati all'approccio simbolico fu l'individuazione di set di istruzioni combinatorie il più complete possibili. Giacché la macchina si limitava ad operare in cieco su oggetti privi di informazione semantica propria, seguendo

semplicemente le istruzioni del programmatore, fu importante che queste istruzioni fossero intrinsecamente intelligenti e performanti. L'Intelligenza Artificiale sarebbe stata tale in quanto algoritmo combinatorio operante su istruzioni intelligenti, prodotto di conoscenza ed esperienza umane. Nascono così i "Sistemi Esperti", programmi IA con istruzioni stilate da esperti di settore, che domineranno il campo per i tre decenni a venire.

Sull'altro ramo della ricerca scientifica, i sistemi subsimbolici concepivano l'intelligenza come qualcosa di non intrinsecamente legata a processi di alta cognizione, operanti su catene di oggetti simbolici e relazioni logiche. L'intelligenza non si costruisce dall'alto, ma emerge dal basso come aggregazione significativa di dati correlati. Un significato, per l'appunto, che non è il prodotto di conoscenza ed esperienza immessa di forza all'interno dei Sistemi Esperti. Esso si trova invece già distribuito all'interno dei dati puri sotto forma di aggregazioni e disgregazioni dinamiche delle fluttuazioni statistiche nelle relazioni tra i dati. Chiariamo, un modo di dire popolare dice: se si giudica un pesce dalla sua abilità di arrampicarsi sugli alberi, lui passerà tutta la sua vita a credersi stupido. E tuttavia giudicheremmo intelligente un

programma artificiale in grado di distinguere la specie con la quale si sta interfacciando sulla base della capacità di questa nel portare a termine questo o quel compito specifico.

Il principio alla base dei sistemi di intelligenza artificiale subsimbolica è esattamente questo: cercare di derivare dai dati aggregati alcune correlazioni e studiare il variare di queste per ricavare una certa legge probabilistica. Dato un certo insieme-campione di pesci è facile notare come ciascun membro del campione possieda alti punteggi nella misurazione delle sue capacità di nuoto e bassi punteggi nella misurazione della sua capacità di arrampicarsi. Ciò è già potenzialmente sufficiente alla macchina per dedurre che, se l'animale con il quale si sta interfacciando ha una sviluppata capacità di nuotare e una ridotta capacità di arrampicarsi, quello che ha davanti è più probabilmente un pesce che una scimmia.

Ciò che la macchina andrebbe in quel caso a intercettare non sarebbero descrizioni simboliche sulla natura dell'oggetto analizzato, ma delle spe-

cifiche e predeterminate caratteristiche quantificabili, delle quali sono studiate le relazioni tra loro nel tempo⁵.

4. Percettrone: il sistema nervoso

La modellizzazione dei neuroni sotto forma di funzioni logiche operanti su output binari avanzata da McCulloch e Pitts fornì un quadro teorico stabile sul quale non si tardò a congetturare: era il 1958 quando lo psicologo americano Frank Rosenblatt avanzava un modello di algoritmo di intelligenza artificiale dichiaratamente ispirato ai neuroni, il percettrone (Rosenblatt 1958). Questo era un algoritmo di apprendimento supervisionato⁶ per la classificazione binaria, ovvero un programma di intelligenza artificiale appositamente

⁵ Si noti come non è strettamente necessario che la quantificazione sia particolarmente graduale: già una possibile ripartizione 0-1 o Sì-No potrebbe permettere alla macchina di discriminare due specie differenti sulla base anche solo di una misurazione. Ad esempio, le domande “vive in mare?” o “può affogare?” potrebbero essere già più che sufficienti per discriminare tra un pesce e una scimmia. Lo studio di più variabili potrebbe fornire ad un programma dotato di abbastanza potere computazionale la capacità di discriminare, ponendo ulteriori domande, tra specie diverse di pesci e specie diverse di scimmie.

⁶ Ovvero un algoritmo addestrato dalla macchina a dedurre dai dati le categorie di appartenenza per poi essere, nel caso, corretto dai programmatori, che forniscono successivamente le categorie appropriate per i dati forniti.

addestrato a fornire risposte binarie a specifiche domande. Forte della modellizzazione di McCulloch e Pitts, Rosenblatt fornì un algoritmo descrittivo del funzionamento dei neuroni tenendo in conto delle tre caratteristiche principali di questi: a) la ricezione di input inviati dall'esterno e la loro somma; b) lo scaricamento da parte del neurone di un output a seconda della forza del segnale complessivo ricevuto; c) la capacità del segnale di output di poter essere eccitatore o inibitore del neurone che lo riceverà in input. Poste queste basi, il modello costruito in analogia da Rosenblatt riusciva a simulare in linea di principio il funzionamento di un neurone biologico in maniera alquanto precisa: sostituito l'input bioelettrico del neurone con un input digitale numerico, posta una funzione di somma sugli input ricevuti e una funzione di segno⁷ sul risultato di quest'ultima, ecco che si otteneva un algoritmo in grado di trasformare vari input binari in un unico output.

Giacché McCulloch e Pitts avevano già dimostrato di poter sfruttare un algoritmo ad output binario di tal fatta per la manipolazione della logica proposizionale, ne risultava immediatamente la

⁷ Per funzione di segno s'intende una funzione che trasforma ogni numero superiore a 0 in 1 e inferiore a 0 in -1.

possibilità di uso per compiti complessi. Spingendosi anche oltre, Rosenblatt dimostrò matematicamente che il suo perceptrone era in grado di distinguere due gruppi di oggetti ogni volta che esisteva un modo chiaro per tracciarne una separazione netta. Non solo, l'intero procedimento non richiedeva l'intervento umano, ma veniva eseguito dal perceptrone modificando autonomamente l'importanza che esso attribuiva ai diversi input ricevuti.

L'algoritmo di Rosenblatt riusciva a partire da una configurazione di avvio completamente casuale fino ad auto-assestarsi su una configurazione ottimale per l'individuazione di una divisione netta tra due classi di oggetti. Dato un insieme casuale di dati e un compito predeterminato, il perceptrone si mostrò in grado di automigliorarsi e usare i dati bruti per dedurre il modo migliore di soddisfare il compito, sintetizzando un comportamento intelligente da informazioni neutre. Il macchinario di Rosenblatt, il Mark I Perceptron, fu utilizzato nel campo del riconoscimento ottico, identificando con successo delle lettere dell'alfabeto attraverso la loro differente forma.

5. Inverno dell'AI: adattamento funzionale

E tuttavia la strada del nuovo e incoraggiante algoritmo matematico non andò molto lontano: nel 1969 Minsky pubblicò un'analisi critica dei limiti del perceptrone di Rosenblatt che minacciò di esservi fatale, mostrandone matematicamente l'incapacità nell'approcciare problemi più complessi⁸ (Minsky, Papert 1969). I circa trent'anni che ne seguirono, periodo noto come "Inverno dell'Intelligenza Artificiale", rappresentarono una fase di declino dell'interesse e del supporto verso la disciplina, sia da parte della comunità scientifica che degli investitori.

Durante questo tempo, la ricerca si concentrò prevalentemente su approcci simbolici e sull'elaborazione dei Sistemi Esperti, limitando significativamente il progresso in altri paradigmi, ma non per questo senza raggiungere importanti successi degni di interesse. Sono questi gli anni di sviluppo di piccole meraviglie dell'Intelligenza Artificiale simbolica come MYCIN, DENDRAL e il robot

⁸ In particolare, il lavoro di Minsky e Papert mostrò che il perceptrone di Rosenblatt e gli algoritmi che lo muovevano non erano in grado di rendere ragione di tutti quei casi assimilabili alla necessità di dividere due o più categorie di oggetti in mancanza di una linea di divisione netta tra le categorie. Questo, ad esempio, il caso dell'operatore logico XOR, rilevante in campo informatico.

Shakey⁹. Il primo algoritmo, sviluppato da Stanford per l'assistenza medica, era in grado di diagnosticare correttamente una malattia infettiva nel 65% dei casi, superando in media i risultati di un medico non specialistico e potendo fornire autonomamente raccomandazioni terapeutiche, nonché dosaggi dei farmaci. Il secondo algoritmo, invece era in grado di costruire autonomamente strutture chimiche coerenti secondo le leggi di composizione al tempo conosciute, scartando quelle improbabili e accelerando in questo modo la scoperta di numerose molecole nel campo della chimica organica e della farmacologia.

Tuttavia, il grande limite dei Sistemi Esperti divenne col tempo sempre più vistoso: essi poggiavano su un'architettura costituzionalmente limitata, giacché tutta la loro efficienza derivava dall'efficienza delle regole di applicazione che venivano loro fornite dagli esperti di settore. Non è difficile immaginare la crescita esponenziale delle regole richieste al programma per continuare il suo lavoro, man mano che i casi da considerare si facevano sempre più specialistici e che gli alberi di

⁹ Il primo progetto a ibridare la robotica, l'Intelligenza Artificiale, la visione digitale e l'elaborazione del linguaggio naturale, andando a costituire un caposaldo delle principali linee di ricerca successive, dalle auto a guida automatica agli LLM di nuova generazione.

ipotesi-deduzione ampliavano il loro ventaglio. L'architettura di funzionamento dei Sistemi Esperti non era infatti in grado di scalare con sufficiente efficienza all'aumentare della complessità dei compiti che le venivano sottoposti. Inoltre, anche le limitazioni tecnologiche del tempo iniziarono a stringere al collo dei programmi, sempre più voraci di potere computazionale.

DENDRAL o MYCIN, benché sorprendenti, erano algoritmi IA che richiedevano estesi periodi di tempo, cosa che ne limitava l'applicazione in contesti reali. Benché la storia dell'Intelligenza Artificiale avesse sorpreso, e neanche durante il lungo inverno si fosse davvero fermata, l'interesse per la materia era sensibilmente calato a causa della frantumazione delle eccessive aspettative e dei proclami ottimistici contro gli scogli dei limiti tecnici e logici che si venivano a incontrare. Anche gli investimenti necessari per i minimi successi andavano oltre le iniziali previsioni. Tuttavia, la ricerca proseguì, seppure con gradualità e in assenza di quel fervore inebriante dei primi anni. Un ramo interessante e fecondo della ricerca IA si spostò ad esempio sul tentativo di sviluppare algoritmi di controllo di posizioni nei giochi a somma

zero¹⁰. La cosa costituì un terreno di ricerca ricco di interesse sia per lo sviluppo di algoritmi di analisi della posizione e ricerca delle mosse migliori, sia per lo sviluppo di algoritmi euristici che permettessero ai programmi di imparare quali alberi di calcolo era sensato intraprendere e quali andavano potati il prima possibile, contribuendo così allo sviluppo di meta-algoritmi di ampio utilizzo.

I risultati non tardarono ad arrivare, portando allo sviluppo di programmi come DeepThought e DeepBlue, che nel 1997 sconfissero Gary Kasparov, campione del mondo di scacchi. Fu quello un turning point importante per l'IA, come per gli scacchi: gli algoritmi di Intelligenza Artificiale smisero di inseguire le performance umane, affrancandosene e ritagliandosi un posto tra i sistemi di cheating, giacché nessun essere umano poteva ormai eguagliarne il livello. Inoltre, la loro applicazione toccava per la prima volta un campo di maggior interesse per un pubblico non specialistico. In realtà poi DeepBlue non comprendeva davvero le regole degli scacchi, ma operava per forza bruta, partendo da un database di aperture ottimali e calcolando fino a duecento milioni di posizioni possibili al secondo. Pura potenza

¹⁰ Ovvero tutti quei giochi dove il guadagno di un giocatore coincide alla perdita dell'altro, come gli scacchi o la dama.

bruta di calcolo, ma tanto bastava per surclassare l'essere umano sul piano della performance.

6. Rinascita

Tuttavia, nelle retrovie della ricerca sul tema qualcosa aveva da tempo iniziato a muoversi: fin dalla stroncatura di Minsky del '69 sulle potenzialità dei perceptron, studi e proposte avevano continuato in sordina ad essere avanzati. Era chiaro a tutti che le barriere computazionali del perceptrone potevano in linea di principio essere superate solo costruendo un'architettura multistrato di perceptron in sequenza e mutualmente dipendenti, ma quello che occorreva era un sistema intelligente che permettesse di legare non solo i perceptron di output del programma a quelli da cui dipendevano, ma anche questi ultimi ai primi.

L'algoritmo di retropropagazione dell'errore, applicato alle reti neurali soltanto all'inizio degli anni Ottanta (Werbos 1982), costituì un fulmine a ciel sereno che avrebbe scosso il mondo della ricerca sulle intelligenze artificiali fin dalle sue fondamenta. L'idea alla base era semplice: inizialmente, calcolare l'errore tra l'output previsto e quello atteso e successivamente "propagarlo all'indietro" attraverso la rete di perceptron, determi-

nando con un'apposita formula quanto ogni perceptrone avesse contribuito alla determinazione dell'errore finale. Il sistema si mostrò da subito estremamente intelligente ed efficiente, perché era in grado di riconoscere l'importanza emergente dei singoli perceptroni nell'economia del calcolo complessivo e di agire in maniera più marcata proprio su di essi. Alternando processi di forward pass (calcolo dagli input agli output) e backward pass (valutazione dagli errori di output agli errori di input), l'intera rete neurale, indipendentemente da quanto grande, otteneva nuove e insperate potenzialità, potendo auto modificarsi in chiave dinamica sotto la spinta dei soli dati di partenza.

Una volta introdotta questa modifica, la ricerca sulle reti neurali profonde (ovvero algoritmi composti da decine e decine di strati di neuroni artificiali in sequenza) si avviò lungo un percorso di successi esponenziali. La rinascita dei sistemi sub-simbolici andò ben oltre le aspettative, surclassando i più farraginosi Sistemi Esperti, grazie alla loro capacità di apprendere direttamente dai dati, senza ricorso a fasi intermedie di programmazione classica per ogni espansione sui dati. In linea di principio, inoltre, i Sistemi Simbolici, programmati sulle conoscenze umane, non avevano la possibilità di andarvi oltre, potendo nella migliore

delle ipotesi costituire un assistente, una sorta di tutor esperto che accompagnasse l'utente fino a livelli di competenze più elevate ma già apprese e formalizzate da esperti e luminari. Le deep neural networks potevano ottenere molto di più grazie alla loro capacità di cogliere, se forniti di un dataset di partenza sufficientemente esteso, correlazioni anche minute, delle quali non era ancora stata scoperta l'esistenza.

Con queste rinnovate potenzialità, la loro applicazione divenne la base per i principali algoritmi di Intelligenza Artificiale, anche quelli non strettamente subsimbolici. Riconoscimento vocale (Hinton *et al.* 2012; Hansen, Hasan 2015), riconoscimento visivo (Krizhevsky *et al.* 2012), generazione di testo (Brown *et al.* 2020), generazione di contenuti audio-visivi (Ramesh *et al.* 2021), applicazioni di diagnostica medica (Miotto *et al.* 2018), è difficile immaginare un campo di applicazione dell'IA non toccato e completamente rinnovato dalle reti neurali profonde e dalla loro capacità di estrapolare relazioni da dati spuri.

Tra i più performanti algoritmi neurali quelli generativi finiranno per avviare un'intera nuova e prolifica branca della ricerca AI. Si tratta di reti neurali addestrate per simulare proprio quel carattere di creatività simil-umana tanto cara ai

partecipanti al workshop di Dartmouth. Sfruttando un addestramento supervisionato dai programmatori – che catalogano i dati forniti – e un apprendimento per rinforzo che premi l’algoritmo quando genera contenuti adatti, si condiziona la rete neurale alla produzione di contenuti sempre più statisticamente coerenti con determinate categorie fornite.

Di largo uso per questi fini sono le CNN (Reti Neurali Convoluzionali, particolarmente performanti nella creazione di immagini), i Trasformatori auto-regressivi, che determinano le probabilità che un determinato elemento sia successivo a tutti quelli precedenti, particolarmente usati per la generazione di testo o i GANs (Reti Generative Avversarie, composte da due algoritmi distinti che cercano di ingannarsi a vicenda in un gioco a somma zero analogo al gioco dell’imitazione di Turing). L’introduzione dei trasformatori, in particolare modo, giocò un ruolo fondamentale per il colpo di reni che ha portato alla diffusione senza limiti degli ultimi anni: l’articolo del 2017 di un gruppo di ricercatori di Google sulle IA introdusse una nuova e stupefacente architettura logica, a base neurale, per consentire ai programmi di interfaccia testuale di tenere traccia del contesto dinamico

in cui le singole parole in un testo si venivano a trovare (Vaswani *et al.* 2017).

Quest'ultimo, imprevisto tassello, fornì la chiave di volta per il superamento delle meccaniche sintattiche che avevano caratterizzato tentativi come quelli del Logic Theorist nella formulazione di discorsi logici, raggiungendo una apparente simulazione della capacità di governare la semantica degli enti logici manipolati, in questo caso le parole. I simil-neuroni diventarono in grado di generare simil-discorsi. Sorprendentemente, il ramo subsimbolico dell'Intelligenza Artificiale si è mostrato il più fecondo dal punto di vista degli applicativi di largo uso. Se i Sistemi Esperti e le IA simboliche, forti delle loro capacità deduttive basate sulla logica e sulle conoscenze solide, costituiscono il campione delle AI di tipo causale, sono state piuttosto le Reti neurali basate sulle correlazioni di dati a raggiungere il trionfo globale. Ma ciò che deve maggiormente far riflettere è come l'esplosione delle potenzialità delle reti neurali coincida con la perdita del controllo da parte del programmatore.

L'algoritmo di retropropagazione consente alla macchina, seppure in maniera fortemente energivora, di apprendere dai dati, indipendentemente dal numero di strati nascosti di neuroni

strettamente interconnessi tra loro. Una volta avviato, l'algoritmo neurale estrapola in maniera forsennata e instancabile correlazioni, per generare una risposta che poi l'ambiente – ovvero il programmatore – giudicherà adatta o meno. C'è in tutto questo una meccanica selettiva di sapore darwiniano che si fa ancor più palese alla luce della rinuncia al controllo del funzionamento interno dell'algoritmo, demandato in maniera pressoché completa ai processi di retro-propagazione al più perfezionabili, ma ormai irrinunciabili.

Le reti neurali, forti di una intrinseca resistenza al rumore di fondo nei dati che vengono loro forniti – complice il loro funzionare per adattamento statistico – costituiscono una “black box” insondabile che può contare miliardi di parametri differenti e codipendenti, scivolando lentamente fuori dalla categoria del semplice strumento passivo, analizzabile e comprensibile, seppure da mente esperta del settore. Ancorate a un complesso arco riflesso di input e output addestrato in modo automatico, dinamico e costante, soggiacciono a due sole logiche ambientali: quella dei dati con i quali scegliamo di alimentarle e quella dell'addestramento cui decidiamo di sottoporle, secondo meccaniche completamente assimilabili al condizionamento operante per rinforzi positivi

e negativi. Tutto ciò sembra suggerire che sarebbe – forse ancor prematuramente – non scorretto iniziare a guardarvi a un intricato, seppur ancora embrionale, sistema artificiale-animale, dotato delle proprie impenetrabili meccaniche interne. Il loro delta di evoluzioni filogenetiche, delle quali molto sommariamente si è riassunto alcuni dei punti principali, ha sperimentato vie multiformi, assestandosi ormai verso una soluzione potente e oscuratamente performante.

L'intelligenza artificiale nella sua declinazione neurale ha diffuso la sua presenza in maniera pervasiva riducendo, semplificando, velocizzando il vecchio, ma anche aprendo ad orizzonti economici e sociali inediti e come bestiame novello sapientemente ibridato promette di cambiare ogni paradigma, adattandosi a noi a tal punto che adattarsi a lei non sarà più questione di scelta, ma di coesistenza.

Bibliografia

Brown T.B. *et al.* (2020), *Language models are few-shot learners*, “arXiv”, <https://arxiv.org/abs/2005.14165>.

Church A. (1936), *An unsolvable problem of elementary number theory*, "American Journal of Mathematics", 58(2), 345-363.

De Fauw J. *et al.* (2018), *Clinically applicable deep learning for diagnosis and referral in retinal disease*, "Nature Medicine", 24(9), 1342-1350.

Goodman J., Cormack G. V., Heckerman, D. (2007), *Spam and the ongoing battle for the inbox*, "Communications of the ACM", 50(2), 24-33.

Grigorescu S., Trasnea B., Cocias T., Macesanu G. (2020), *A survey of deep learning techniques for autonomous driving*, "Journal of Field Robotics", 37(3), 362-386.

Hansen J.H.L., Hasan T. (2015), *Speaker Recognition by Machines and Humans: A tutorial review*, "IEEE Signal Processing Magazine", 32(6), 74-99.

Hinton G. *et al.* (2012), *Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups*, "IEEE Signal Processing Magazine", 29(6), 82-97.

Hoy M.B. (2018), *Alexa, Siri, Cortana, and Google Assistant: A comparison of speech-based natural user interfaces*, "Medical Reference Services Quarterly", 37(1), 81-88.

Krizhevsky A., Sutskever I., Hinton, G.E. (2012), *ImageNet classification with deep convolutional*

neural networks, “Advances in neural information processing systems”, 1097-1105.

McCarthy J. (2004), *What is artificial intelligence?*, Stanford University, Stanford.

McCarthy J., Minsky M.L., Rochester N., Shannon C.E. (1955), *A proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, in “AI Magazine”, 2016, 27(4).

McCulloch W.S., Pitt W. (1943), *A logical calculus of the ideas immanent in nervous activity*, “The Bulletin of Mathematical Biophysics”, 5(4), 115-133.

Minsky M., Papert S. (1969), *Perceptrons: An introduction to computational geometry*, MIT Press, Cambridge MA.

Miotto R., Wang F., Wang S., Jiang X., Dudley J.T. (2018), *Deep learning for healthcare: review, opportunities and challenges*, “Brief Bioinform”, 19(6), 1236-1246.

Nilsson N. J. (2009). *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge University Press, Cambridge MA.

Parkhi O.M., Vedaldi A., Zisserman A. (2015), *Deep face recognition*, Proceedings of the British Machine Vision Conference (BMVC).

Phua C., Lee V., Smith K., Gayler R. (2010), *A comprehensive survey of data mining-based fraud detection research*, "Artificial Intelligence Review", 34(4), 341-381.

Ramesh A. *et al.* (2021), *Zero-Shot Text-to-Image Generation*, <https://arxiv.org/abs/2102.120-92>.

Rosenblatt F. (1958), *The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain*, "Psychological Review", 65(6), 386-408.

Simon H.A. (1965), *The Shape of Automation for Men and Management*, Harper & Row, New York.

Smith B., Linden G. (2017), *Two decades of recommender systems at Amazon.com*, "IEEE Internet Computing", 21(3), 12-18.

Turing A.M. (1936), *On computable numbers, with an application to the Entscheidungsproblem*, "Proceedings of the London Mathematical Society", 2(1), 230-265.

Vaswani A. *et al.* (2017), *Attention Is All You Need*, "Advances in Neural Information Processing Systems (NeurIPS)", 5998-6008.

Vinge V. (1993), *Technological Singularity*, Conference paper, NASA, Lewis Research Center.

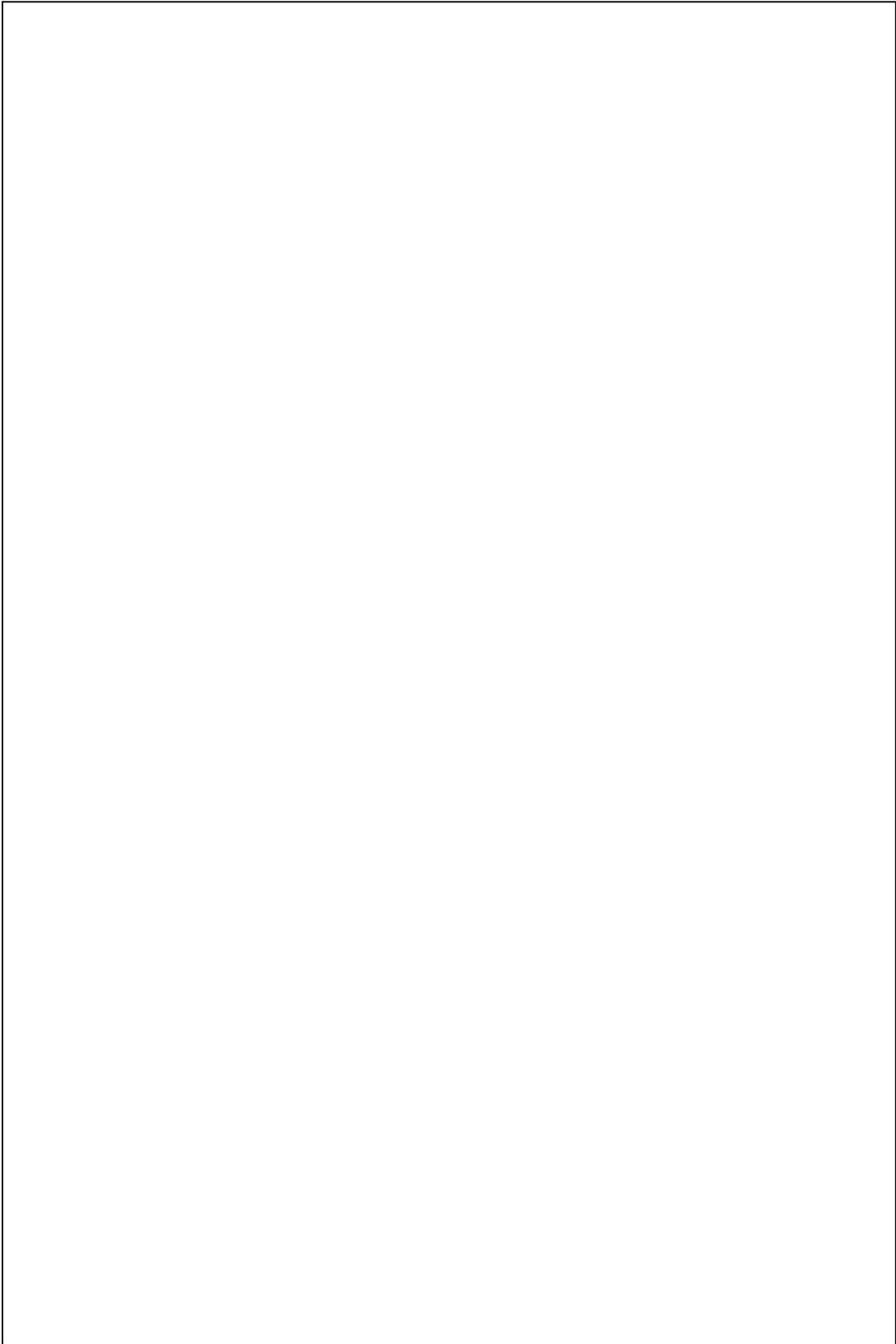
Von Neumann J. (1945), *First Draft of a Report on the EDVAC*, Moore School, University of Pennsylvania, Philadelphia (reprinted in 1993, "IEEE Annals on the History of Computing", 15(4), 27-43).

Werbos P.J. (1982), *Applications of advances in nonlinear sensitivity analysis*, in Drenick R., Kozin F. (a cura di), *System Modeling and Optimization: Proc. IFIP*, Springer, Berlin.

**Centro
Editoriale
Accademico
unicollege**



Lorenzo Grifone Baglioni è autore del progetto grafico editoriale della collana.



unicollege
working
papers
unicollege
working
papers
unicollege
working
papers
unicollege
working
papers
unicollege