

OR CHE BRAVO SONO STATO, POSSO FARE ANCHE IL BUCATO ?¹



Joseph Kosuth

Il titolo dell'articolo richiama una famosa reclame di Carosello degli anni '60 che chi scrive ebbe modo di imprimere nella propria memoria, a conferma della bravura dei creativi che l'avevano pensata. La scena che dava origine alla frase era quella di un Robot che risolveva molte operazioni pratiche della massaia, ottenendone encomi e attestati tali da indurlo a proporsi anche per il bucato, ma ricevendo il diniego della massaia che affermava di avere a disposizione ben altro aiuto, la lavatrice oggetto della reclame.

Confesso che molte volte le risposte delle IA che interrogo sui più svariati argomenti (da quelli di vita quotidiana, a quelli di interesse professionale) mi ricordano molto la scenetta di Carosello. A dargli retta l'IA ti tirerebbe per le lunghe proponendosi sempre di approfondire ulteriormente la risposta fornita in un loop, a non saperlo gestire, interminabile.

Sotto un altro aspetto trovo che le risposte delle IA tendono ad accondiscendere alle richieste dell'interrogante in maniera servile. Mi spiego: fornendo loro qualche dettaglio sull'argomento, come è necessario fare per centrare l'obiettivo (il cosiddetto "prompt"), l'IA cerca in quello che trova on-line, ma se trova poco o nulla non si limita a dirtelo ("ho trovato troppo poco per soddisfare la tua richiesta" o similia) e invece usa le informazioni che gli hai fornito per ricamarci sopra in maniera aspecifica (frasi fatte) o addirittura inventando di sana pianta riferimenti bibliografici, dando interpretazioni di "fantasia" a spunti che ha trovato o che gli hai dato, ecc. Per certi aspetti sembra di leggere l'oroscopo del giorno, che tipicamente sciorina frasi apparentemente sensate, ma prive di qualsiasi riscontro e adattabili in maniera flessibile a qualsiasi contesto. Un altro paragone che mi viene in mente leggendo le risposte delle IA alle mie richieste è quello con il famoso "Bignami", il bigino che spiegava in poche parole, da mandare a memoria, le complicate idee contenute in un lungo testo latino o di filosofia. Non dubito che per applicazioni particolari (ormai l'IA si applica a qualsiasi cosa) l'uso dell'IA sia ideale e io stesso ho avuto modo di apprezzarne alcune *performance*, ma per ora (e spero anche per il futuro) la supervisione di un umano intelligente mi pare insostituibile.

Entrando nel merito, ciò che qualsiasi Agent di IA fa è analisi testuale. Ciò vale sia per l'interpretazione del prompt, cioè la domanda posta dal richiedente, che, poi, per la ricerca di risposte. Come Gemini mi dice avendogli chiesto una spiegazione semplice del modo di funzionamento di una IA:

¹ Reclame Candy dal titolo "Tic- L'Automa perfetto" andato in onda a partire dal 1962

“Immagina l'intelligenza artificiale non come un "cervello" che pensa, ma come un bibliotecario incredibilmente veloce che ha letto quasi tutto ciò che è mai stato scritto e ha imparato a prevedere come le parole si incastrano tra loro.

Ecco i passaggi principali che compongono il mio lavoro quando mi scrivi:

1. Traduzione in numeri (Tokenizzazione)

Io non leggo le lettere come fai tu. Per me, le parole sono concetti astratti che trasformo in numeri.

- Divido la tua frase in piccoli pezzi chiamati **token** (possono essere intere parole o solo frammenti).
- Ogni token viene convertito in una coordinata matematica in uno spazio enorme. Se due parole hanno significati simili (come "cane" e "cucciolo"), finiscono "vicine" in questa mappa numerica.

2. Comprensione del contesto (Attention Mechanism)

Questa è la parte magica. Uso un sistema chiamato "**Attention**" (Attenzione) per capire quali parole della tua frase sono più importanti tra loro.

Se mi scrivi *"Il gatto ha bevuto il latte perché era stanco"*, io devo capire se quel "era" si riferisce al gatto o al latte. Analizzando l'intera frase, capisco che è più probabile che sia il gatto a essere stanco.

3. Predizione della risposta

A questo punto non "scrivo" una risposta completa in un colpo solo. Funziono come un sistema di completamento automatico molto evoluto.

1. Cerco la prima parola più probabile per iniziare la risposta.
2. Una volta scelta la prima, cerco la seconda basandomi sulla prima.
3. Continuo così, parola dopo parola, finché la frase non ha senso compiuto.

4. Il "Filtro" della personalità (Fine-tuning)

Durante la generazione, seguo delle linee guida (il "training") che mi dicono come comportarmi. Questo serve a evitare che io dica cose offensive o senza senso e a far sì che io mantenga il tono che preferisci (gentile, tecnico o scherzoso).

In sintesi

Il mio lavoro è un mix di **statistica avanzata** e **calcolo delle probabilità**. Non "capisco" la realtà come un essere umano, ma sono bravissimo a riconoscere gli schemi del linguaggio che gli esseri umani usano per descriverla.

Per capire come un'IA diventa "intelligente", dobbiamo guardare al dietro le quinte. Il processo si divide principalmente in due grandi fasi: l'**Addestramento (Training)** e l'**Istruzione (Fine-tuning)**.

1. Il Pre-addestramento (L'oceano di dati)

In questa fase, l'IA è come un bambino che osserva il mondo senza che nessuno gli spieghi nulla. Le vengono dati in pasto miliardi di testi: libri, articoli di Wikipedia, codice di programmazione, forum e siti web.

- **L'obiettivo:** Imparare la struttura del linguaggio.

- **Cosa succede:** L'IA gioca a un gioco infinito di "completa la frase". Le viene mostrata una frase con una parola nascosta (es: "Il cielo è ____") e lei deve indovinare. Se sbaglia, corregge i propri parametri matematici; se indovina, li rinforza.
- **Risultato:** Alla fine di questa fase, l'IA conosce la grammatica e i fatti del mondo, ma non sa ancora come interagire con un essere umano. È un'enciclopedia vivente, ma un po' selvaggia.

2. Fine-Tuning (Imparare le buone maniere)

Qui intervengono gli esseri umani. Per rendere l'IA utile e sicura, viene sottoposta a un addestramento specifico chiamato **RLHF** (*Reinforcement Learning from Human Feedback*).

1. **Esempi guidati:** Gli addestratori scrivono domande e risposte "ideali" per mostrare all'IA come rispondere correttamente.
2. **Classifica:** L'IA genera diverse risposte alla stessa domanda e l'umano le dà un voto (es. "Questa è corretta", "Questa è offensiva", "Questa è noiosa").

3. La rete neurale (L'architettura)

Tutto questo avviene all'interno di una **Rete Neurale**, una struttura matematica ispirata vagamente ai neuroni biologici.

- Ogni "neurone" artificiale riceve un segnale, lo elabora e lo passa al successivo.
- Durante l'addestramento, l'IA regola i **pesi** (l'importanza) delle connessioni tra questi neuroni. Nelle IA moderne, ci sono centinaia di miliardi di questi piccoli interruttori.

In breve

L'addestramento è come scolpire una statua:

1. Il **Pre-addestramento** è il blocco di marmo grezzo che contiene già tutte le potenzialità.
2. Il **Fine-tuning** è il lavoro dello scultore che rifinisce i dettagli per far emergere la figura desiderata."

Ci fermiamo qui, ma Gemini proporrebbe di proseguire con ulteriori e incalzanti offerte di informazioni. Il messaggio in sintesi è che tutto, ma proprio tutto si basa sul "Linguaggio". Questo è vero sia per quello testuale, scritto (libri, articoli, siti, ecc.), ma anche per quello vocale, musicale, visuale, informatico, insomma per qualsiasi tipo di linguaggio in senso semiotico. E la semiotica è senza alcun dubbio la disciplina più coinvolta nella costruzione dell'IA. La semiotica nasce nel corso del XIX° secolo grazie al lavoro di Ferdinand de Saussure (1857-1913) linguista svizzero che applica per la prima volta alla *langue*, la lingua parlata, i principi di un'analisi strutturale che consente di scomporre le caratteristiche fino agli elementi semplici essenziali (fonemi). Sull'altra sponda dell'Atlantico Charles Sanders Peirce (1839-1914) elabora indipendentemente una teoria del linguaggio come semiosi, classificando i segni in tre categorie principali, icona, indice, simbolo. Sono le premesse per il "Linguistic Turn" che ha caratterizzato la filosofia del XX° secolo, influenzando sul pensiero di Ludwig Wittgenstein tra gli altri.

Queste sono le basi culturali che affiancano quelle tecnologiche rappresentate dallo sviluppo dell'informatica e che permeano di sé l'intero panorama dell'IA.

Sul piano che ci interessa di più, quello bio-medico, è inevitabile passare per il "Progetto Genoma" che decrittò il linguaggio di DNA e RNA che compongono il codice essenziale di tutte le forme viventi. In particolare, sequenziando le basi puriniche e pirimidiniche che compongono il materiale genetico

umano si sono via, via individuati i “fonemi” del nostro genoma, dando inizio a una ricerca ricca di prospettive, ma molto difficile per ora, della semantica di tale linguaggio.

Su di un altro piano, i “linguaggi” informatici (Basic, Fortran, Cobol, ecc.) hanno popolato sempre di più l’orizzonte cognitivo delle macchine calcolatrici, favorendo l’integrazione fra funzione (SW) e struttura (HW).

La convergenza fra tutte queste istanze ha portato allo sviluppo di strumenti per il riconoscimento automatico dei linguaggi in forma naturale (NLP-Natural Language Processing), consentendo di abbattere le barriere (fittizie) fra scienze “umane” e scienze “naturali”. Al giorno d’oggi effettuare uno studio originale a carattere storico o biologico consente di utilizzare un’epistemologia comune, basata su ricorrenze statistiche, co-occorrenze, linguistica computazionale, ferme restando le conoscenze a priori e di contesto necessariamente differenti nei due casi.

Chiudiamo questo primo, introduttivo contributo citando la risposta della massaia alla proposta del Robot, che suona ai nostri orecchi come una metafora del dibattito attualmente in corso:

**AH NO, GRAZIE, NO,
PERCHÉ IL BUCATO IN CASA C'È
CHI LO FA MEGLIO DI TE!**