

Storia dell'IA con una piccola nomenclatura

di Carlo Bonotto

“Un tale voleva arrivare a conoscere la mente, non in natura, bensì in un suo grande calcolatore personale. Gli chiese [...] “Calcoli che penserai mai come un essere umano?”. La macchina allora si mise al lavoro per analizzare le proprie abitudini di calcolo; infine stampò la risposta su un foglio di carta [...]. L'uomo corse a vedere la risposta e trovò, nitidamente stampate, le seguenti parole: “QUESTO MI RICORDA UNA STORIA”. [G. Bateson *Mente e natura* ed Adelphi - introduzione pagg. 27-28]

Nella citazione di Bateson, in cui c'è anche un'importante sottolineatura sulla caratteristica tipica dell'intelligenza umana di “pensare per storie”, c'è un riferimento sottinteso al cosiddetto “**Test di Turing**”, che lo scienziato britannico enunciò nel 1950 e che è noto anche col titolo “**Imitation game**”.

Il titolo si riferisce ad un gioco molto in voga nell'Inghilterra vittoriana, di cui Turing propose una variante: due stanze chiuse contengono rispettivamente una macchina e un essere umano. Un intervistatore interroga per iscritto chi si trova al di là del muro, per capire chi è dei due sia l'essere umano. Se la macchina riesce a ingannare l'intervistatore, supera il Test e dimostra un *comportamento* intelligente, che *simula* il pensiero umano.

Nel 1956 l'informatico statunitense John Mc Carthy, durante un seminario estivo di un college del New Hampshire, coniò il termine di **IA (Intelligenza Artificiale)** – in Inglese AI - per indicare un nuovo filone di studi rivolti a riprodurre tramite programmi informatici le caratteristiche dell'intelligenza umana. Molti degli studi prodotti nell'ambito della cibernetica sono confluiti in questo settore di ricerca. Anzitutto va sottolineato il fatto che il campo di ricerca attorno all'IA si presenta fin dagli inizi come un campo **interdisciplinare**. Oltre agli informatici sono presenti filosofi, psicologi, linguisti, neuroscienziati, ingegneri cibernetici ecc. Troviamo nomi come Shannon, Minski, von Neuman, Searle, solo per fare qualche esempio.

All'inizio il dibattito si focalizza attorno a quella che, più tardi, sarà definita **IA forte** (o **estesa**).

In altri termini la ricerca si focalizza attorno al progetto di riprodurre artificialmente, tramite strumenti informatici, il funzionamento dell'intelligenza umana, anche per rispondere all'istanza di Turing. Il dibattito è ulteriormente complicato dal fatto che **manca una definizione completa e condivisa di cosa sia l'intelligenza**. La risposta è diversa, a seconda delle varie scuole psicologiche e filosofiche.

Per non dilungarmi troppo, in questa sede mi limiterò ad una esposizione semplificata:

Secondo me, in questo dibattito emergono, fondamentalmente, due posizioni fortemente alternative:

Da un lato chi pensa che la caratteristica fondamentale dell'intelligenza sia basata sul *calcolo logico – razionale* e pensa che questa capacità possa essere tradotta in un programma informatico. In questa visione la mente razionale sarebbe il software e il cervello l'hardware. *Sarebbe quindi possibile sviluppare un software che riproduca l'intelligenza utilizzando l'hardware del computer.*

Dal lato opposto si colloca chi, oltre a sottolineare la complessità della mente umana, difficilmente riducibile ad una serie di algoritmi, sottolinea il *fondamento biologico dell'Intelligenza, che non può essere semplicemente trasferita su un supporto di silicio* (anche e soprattutto per la sua inestricabile connessione con la sfera emotiva). Tra i sostenitori di questa posizione si colloca ovviamente Damasio.

Prima di proseguire, è forse opportuno precisare il significato di **Algoritmo**. Vi propongo una definizione ricavata da Wikipedia: “L'algoritmo è un **elenco d'istruzioni dettagliate, elaborate per svolgere una determinata attività o risolvere un problema specifico**”. Può presentarsi come “schema o procedimento sistematico di calcolo” o, più semplicemente, sotto forma di ricetta di cucina. Ovviamente gli algoritmi usati in informatica debbono essere codificati in maniera comprensibile al computer.

Tornando al dibattito sulla **IA generale** o **forte**, la situazione si trascina per anni, anche perché la psicologia e le neuroscienze, anziché fornire una soluzione definitiva, mostrano la complessità crescente del problema. Qui accennerò solo alla teoria di Gardner sulle intelligenze multiple. *Per Gardner non esiste un solo tipo di intelligenza, ma almeno nove: logico-matematica, linguistica, musicale, spaziale, corporeo-cinestetica, interpersonale, intrapersonale, naturalistica, esistenziale.*

Mentre la prospettiva dell'IA forte viene progressivamente a collocarsi in un punto morto si sviluppano una serie di ricerche incentrate sull' **IA debole** (o **ristretta**). *Non si tratta più di riprodurre l'intelligenza nella sua complessità, ma solo alcuni comportamenti, che se eseguiti da esseri umani, sono comunemente considerati intelligenti.* Non è considerato fondamentale che determinati compiti considerati intelligenti vengano eseguiti secondo i metodi tipici degli esseri umani, ma *si utilizzano strategie funzionali al computer per ottenere i risultati voluti.*

Si fa notare che l'uomo ha costruito strumenti in grado di volare, ma che questi, nel loro funzionamento, si differenziano dal volo degli uccelli o degli altri animali volanti: gli aerei non sbattono le ali.

Prevale dunque una visione di tipo pragmatico e, spesso, decisamente comportamentista.

Negli anni sono state realizzate molte soluzioni legate alla robotica come automi semoventi con ampia "autonomia" (vedi Curiosity su Marte), soluzioni automatiche nell'ambito dell'industria, programmi di ricerca avanzata come Google, assistenti vocali e "navigatori".

La caratteristica comune di questi programmi, che li rende in una certa misura "intelligenti", è la *capacità di apprendere dagli errori* o, comunque, di poter essere facilmente modificati in caso di cambiamenti del contesto o di nuove acquisizioni di dati.

Un filone importante nella ricerca rappresentano i programmi che utilizzano le cosiddette "**reti neurali**", un sistema di hardware che imita le reti neurali del nostro cervello (inventate già negli anni 40 da Mc Culloch e Spits che poi confluiranno nel gruppo dei Cibernetici), permettendo operazioni in parallelo e non solo rigidamente sequenziali. *Questi programmi sembrano essere molto promettenti sul fronte dell'apprendimento automatico (**machine learning**), anche se la complessità delle reti utilizzate è infinitamente inferiore a quella del nostro cervello.*

Prima di esaminare le applicazioni di ultimissima generazione, farò una breve ricapitolazione sui programmi che, negli ultimi 30 anni hanno segnato un traguardo nell'applicazione dell'IA percepibile anche ai non addetti ai lavori:

- 1996 Il computer Deep Blue, sviluppato dall'IBM, sconfigge il campione mondiale di scacchi Garry Kasparov (fino a pochi anni prima molti esperti negavano una simile possibilità)
- 2011 Il programma Watson della IBM sconfigge i campioni del quiz televisivo Jeopardy, molto seguito negli USA, grazie a prestazioni che, per alcuni esperti equivarrebbero al superamento del test di Turing.
- 2016 Il programma Alphago, addestrato sulle reti neurali grazie al **deep learning** (apprendimento profondo, basato su un sistema di reti neurali disposte su più strati), riesce a sconfiggere il più forte giocatore al mondo di GO, gioco diffuso in Estremo Oriente di incredibile complessità (sembra che ci siano più posizioni possibili in questo gioco che atomi nell'universo).

Oltre alla tecnologia delle reti neurali e al continuo aumento di potenza di calcolo delle macchine un dato fondamentale dei continui progressi è stato, negli ultimi anni, la possibilità di accedere ad enormi quantità di dati (**big data**) tramite Internet. [vedi *Intelligenza artificiale* ed. Le frontiere della scienza National Geographic, Milano 17 gennaio 2018 pagg. 61-64].

A questo punto non possiamo non illustrare le ultime applicazioni dell'IA, che negli ultimissimi anni sono state spesso al centro delle cronache e hanno messo a disposizione del pubblico strumenti in grado di operare con "relativa" facilità: le cosiddette **IA generative**. Queste sono in grado di realizzare prodotti tradizionalmente considerati "intellettuali" come testi, immagini, suoni e musica.

Ricordo l'ormai famoso Chatgpt, per i testi, e Midjourney, per la grafica. I problemi relativi all'uso di questi programmi sono diversi: dalla riproduzione degli stereotipi più diffusi nella Rete, alla possibilità di riprodurre "falsi" non riconoscibili dai non esperti, alle cosiddette "allucinazioni", fino ai problemi occupazionali che si potrebbero determinare con la sostituzione di operatori umani con programmi informatici più efficienti.

Qui mi limito a sottolineare i nuovi problemi che l'uso di questi strumenti da parte degli studenti pone all'insegnante, ma anche le opportunità che possono offrire all'insegnamento, laddove si vogliano

guidare gli alunni a ricercare in maniera critica e a porre le domande “giuste” per ottenere risultati soddisfacenti. “[Lo studente] non deve imparare dei pensieri, ma a pensare.”(E. Kant). [La citazione è riportata nel volume di AA.VV. *Scuola e Intelligenza artificiale* Carocci editore]

- **Problemi etici**

Floridi, autore italiano di livello internazionale, ha individuato 5 direttive etiche secondo le quali andrebbe orientata la produzione ed esercitato il controllo nel settore: *beneficenza* (orientamento verso il bene), *non maleficenza* (contro gli usi impropri o illegali), *autonomia* (rispetto alla macchina), *giustizia* (contro le discriminazioni), *esplicabilità* (intelligibilità, ma anche **responsabilità**). [L. Floridi *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide* Raffaello Cortina Editore Milano 2022]

E' da poco stata emanata una direttiva europea, ritenuta dagli esperti di grande rilevanza, ma di cui al momento si sa poco, al di fuori degli addetti ai lavori, mentre sono disponibili online le linee guida della Comunità europea per un'educazione all'IA [Ethical guidelines on the Use of Artificial Intelligence (AI) and data in Teaching and Learning for Educators] reperibili online.

Uno dei temi di attualità riguardo all'uso delle IA generative è quello legato agli **stereotipi socialmente diffusi (bias cognitivi)** come quelli di genere, razziali, classisti, ecc., che peraltro precedono l'uso degli strumenti generativi.

Per concludere, non posso fare a meno di ricordare che molti degli interrogativi che oggi ci poniamo sull'IA o su robot autonomi guidati dall'IA sono stati anticipati dalla letteratura e dal cinema di fantascienza (pensiamo ad Hall 9000 o agli androidi di Blade runner).

In particolare, Isac Asimov già negli anni '40 sentiva il bisogno di porre dei limiti “etici” all'azione dei suoi Robot attraverso l'uso delle famose *3 leggi della robotica*:

- 1) Un robot non può arrecare danno a un essere umano né può permettere che, a causa del proprio mancato intervento, un umano subisca alcun danno.
- 2) Un robot deve obbedire agli ordini degli esseri umani, a meno che tali ordini non siano in conflitto con la prima legge.
- 3) Un robot deve proteggere la propria esistenza a patto che questa non contrasti con la prima o la seconda legge.

Questi riferimenti, ci possono ancora tornare utili? A voi la risposta.

Vacanza – studio del Circolo Bateson a Vicoforte 25-30 Agosto 2024