

L'intelligenza di ALAN TURING

Se ogni strumento riuscisse a compiere la sua funzione o dietro un comando o prevedendolo in anticipo e, come dicono facciano le statue di Dedalo o i tripodi di Efesto i quali, a sentire il poeta, "entran di proprio impulso nel consesso divino", così anche le spole tessessero da sé e i plettri toccassero la cetra, i capi artigiani non avrebbero davvero bisogno di subordinati, né i padroni di schiavi.

ARISTOTELE, *La Politica*

Il 31 agosto del 1955 quattro personaggi, geniali e visionari, redigono un documento, la cosiddetta "proposta di Dartmouth", che prevede la convocazione di una conferenza dedicata a

uno studio approfondito sull'intelligenza artificiale, da parte di dieci esperti per la durata di due mesi, da condursi durante l'estate del 1956 presso il Dartmouth College ad Hanover, nel New Hampshire.

Scopo dello studio è:

procedere sulla base della congettura che ogni aspetto dell'apprendimento o di qualsiasi altra caratteristica dell'intelligenza possa essere descritto in maniera così precisa da poter essere simulato da una macchina.

I personaggi in questione sono John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon.

Marvin Minsky, che nel 2005 prenderà parte anche alla celebrazione del cinquantenario della conferenza di Dartmouth, dichiarerà la sua ferma convinzione sulle potenzialità dell'intelligenza artificiale affermando che

una volta raggiunto il livello medio dell'intelligenza umana, la macchina potrà cominciare a educare sé stessa. Entro pochi mesi (da quel momento) arriverà al livello di genio. E dopo qualche altro mese, il suo potere sarà enorme.

Al lavoro di Dartmouth non poté partecipare colui che più di ogni altro aveva creduto nella possibilità di dar vita ad un'altra forma di intelligenza: Alan Turing. Matematico, pioniere dell'informatica, durante la seconda guerra mondiale aveva combattuto e vinto la sua personale battaglia contro Enigma, il dispositivo elettromeccanico utilizzato dalle forze armate tedesche per cifrare e decifrare messaggi. Si trattò probabilmente di una battaglia decisiva per le sorti della guerra.

La vita di Alan Turing non fu serena né facile. Alle difficoltà ad accreditarsi presso l'apparato militare inglese, seguì, a guerra finita, la condanna da parte

di una società intollerante nei confronti di un diverso. Sconvolto da una crudele terapia imposta dal tribunale, si suicidò, si dice, con una mela avvelenata.

E' a Turing che va attribuita l'ideazione dello schema di principio su cui si basano tutti gli elaboratori elettronici, dal primordiale ENIAC ai più moderni personal computer.

Con Turing si assiste ad un cambio di paradigma: dalla macchina passiva esecutrice di calcoli ed operazioni logiche al dispositivo che un giorno sarà in grado di fornire risposte per le quali l'uomo fa ricorso all'intelligenza.

Nel diciannovesimo secolo il matematico Charles Babbage aveva invano sognato di costruire una *Macchina Analitica*, coinvolgendo nell'impresa Ada Lovelace, figlia di Lord Byron. Gli sforzi compiuti da entrambi non condussero al risultato atteso per i limiti della tecnologia dell'epoca e per lo scetticismo del governo inglese.

Le prestazioni attese della macchina vennero così descritte da Ada Lovelace:

la macchina analitica non ha la pretesa di creare nulla. Non può fare altro che quello che le ordiniamo di eseguire. Può analizzare, ma non ha le capacità di anticipare le connessioni o le verità analitiche. L'unica sua funzione è quella di rendere accessibili le conoscenze acquisite.

Con riferimento a questa affermazione gli storici di informatica parlano di "regime Lovelace".

Dopo quasi un secolo Turing pose le basi per la caduta di questo regime, nel concepire una forma di intelligenza diversa da quella umana e dalle potenzialità addirittura superiori.

Nel 1950 Alan Turing aveva scritto un articolo, quanto mai profetico e provocatorio, sull'intelligenza artificiale. Era intitolato "Computer Machinery and Intelligence" e apparve sulla rivista "Mind". L'articolo comincia con la frase:

Mi propongo di considerare la domanda: "Le macchine possono pensare?"

Poiché, come egli sottolinea, questi sono termini pregnanti, si dovrebbe ovviamente cercare un modo operativo di affrontare la domanda; e lo si può trovare, egli suggerisce, in ciò che chiama "il gioco dell'imitazione", oggi conosciuto come *test di Turing*.

Nel proporre il suo famoso test, atto a giudicare il livello di intelligenza di una macchina, Turing traccia una analogia con il gioco che lui stesso ha chiamato "uomo-donna".

Questo gioco si può eseguire con tre persone: un esaminatore in una stanza, e un uomo e una donna in un'altra. Scopo del test è fare in modo che l'esaminatore scopra qual è l'uomo e qual è la donna. Entrambi i giocatori

possono raccontar bugie, o dare risposte false, a domande del tipo: "E' femmina?" e : "Ha i capelli lunghi?". Turing sostiene che se i giocatori maschio e femmina sono abbastanza intelligenti, l'esaminatore non riuscirà a indicare il loro sesso con la certezza di averci azzeccato.

Turing è convinto che un giorno la stessa impasse si raggiungerà se uno dei giocatori sarà sostituito da una macchina.

Molti hanno raccolto la sfida di Turing mostrando scetticismo nei confronti del modello computazionale della mente su cui è basata l'intelligenza artificiale. Secondo costoro infatti tale modello non prenderebbe in considerazione gli aspetti fondamentali della mente umana, che sono la **coscienza** e l'**intenzionalità**.

Il filosofo americano John Searle ha formulato in più occasioni critiche nei confronti della cosiddetta *intelligenza artificiale forte*, quella che non si limita a considerare il computer come un utile strumento di indagine della mente umana, ma si spinge ad affermare che, con opportuni programmi, esso diviene analogo alla mente umana ed è quindi capace di comprendere e di avere altri stati cognitivi. La sua posizione è chiaramente espressa in un esperimento mentale da lui ideato, quello della cosiddetta "stanza cinese". Questo esperimento immaginario ha lo scopo di mostrare che la semplice manipolazione sintattica di simboli formali non costituisce di per sé una **semantica**, ovvero non garantisce che il sistema comprenda ciò che sta facendo. Più in generale, l'esperimento pone alla nostra attenzione il fatto che un sistema che si comporta come se avesse **stati mentali** non ci permette di essere sicuri che esso possieda realmente questi stati.

Si immagini di chiudere in una stanza una persona che non conosce una parola di cinese. La persona ha a disposizione due gruppi di fogli: sui fogli del primo si trova una serie di caratteri cinesi, sugli altri fogli ci sono delle istruzioni su come utilizzare i caratteri stessi.

Il compito assegnato alla persona in questione è di produrre degli insiemi di caratteri cinesi (risposte), seguendo unicamente le istruzioni ricevute, ogni volta che riceve dall'esterno degli insiemi di caratteri cinesi (domande).

Il punto fondamentale dell'esperimento è che a un cinese che ponga le domande e legga le risposte ricevute, la persona chiusa nella stanza appare come se fosse in grado di comprendere il cinese, mentre, in realtà, si limita a manipolare simboli senza significato sulla base di istruzioni.

Per Searle e per gli altri critici dell'intelligenza artificiale "forte" la mente è qualcosa di più della semplice manipolazione di simboli formali: essa ha dei contenuti. Quando pensiamo a qualcosa, i termini e i concetti adoperati non rappresentano soltanto simboli formali, ma possiedono per noi un significato, cioè una semantica.

A obiezioni di questo genere Turing risponde argomentando che, via via che la complessità delle macchine accrescerà, potremo avere delle sorprese e fa un parallelo con il reattore a fissione; sotto una certa dimensione "critica" non accade gran che, ma, superata la dimensione critica, cominciano a sprizzare

faville. Lo stesso, forse, accade per i cervelli e per le macchine, la maggior parte dei quali sono oggi "subcritici".

Ma il fronte degli oppositori rispetto alla sfida di Turing annovera anche coloro che formulano e paventano inquietanti scenari futuri. Se l'intelligenza artificiale può migliorarsi, è possibile che essa pervenga a una coscienza di sé tale da consentirle di soppiantare, in una sorta di processo evolutivo, la stessa intelligenza biologica che l'ha creata?

L'ipotesi che l'umanità possa essere sostituita da computer superintelligenti è sembrata a molti, anche tra tra i massimi esperti di informatica, un'idea funesta.

Joseph Weizenbaum, professore di computer science, è l'autore del famoso programma ELIZA, che simula uno psicoterapeuta con cui è possibile avere una conversazione terapeutica e intima (servendosi del terminale di un calcolatore) e che è spesso citato come il più straordinario esempio reale di un calcolatore che "supera" il test di Turing. Ebbene, Weizenbaum stesso è spaventato dalle prospettive dell'intelligenza artificiale che giudica ripugnanti e disgustose per ogni persona civile.

La letteratura fantascientifica ha generosamente trattato l'argomento, con toni drammatici, catastrofici e, a volte, con un briciolo di ironia.

Frederic Brown, in un suo famoso racconto (*Answer*, 1954), descrive la costruzione della prima supermacchina onnisciente. Il suo progettista le chiede: "Dio esiste?". E la macchina, dopo aver verificato di avere sotto controllo l'alimentazione elettrica dei propri circuiti, risponde con voce tonante: "Sì, **ora** esiste!". L'inorridito progettista si precipita a disinserire la spina; ma la macchina lo precede e lo colpisce a morte.

Sandro Taglienti, Associazione I SETTE