

SCENARI

Ma se è generale l'intelligenza artificiale pensa?

La nuova frontiera della tecnologia è sviluppare macchine capaci non solo di svolgere compiti definiti, ma di applicare schemi di risoluzione a qualsiasi situazione. Come fa l'uomo. Prima di parlare di intelligenza artificiale generale e di singolarità dobbiamo tornare al 1950, quando Alan Turing si chiese se «le macchine possono pensare». Quella domanda rappresentò uno spartiacque concettuale che separò le ricerche svolte fino ad allora per capire il funzionamento del cervello umano, per costruire una macchina in grado di decrittare i messaggi scambiati dall'esercito nazista, per progettare e realizzare la bomba atomica che distrusse due città giapponesi, da quelle successive finalizzate a costruire macchine pensanti. La domanda è collegata al test che Turing propose non per dichiarare che una macchina pensa, ma per osservare che si comporta come tale. La stessa domanda può essere interpretata diversamente, considerando le due diverse accezioni del verbo "potere": essere in grado di fare qualcosa, essere autorizzato a fare qualcosa. Nel primo caso sappiamo con certezza che Turing si concentrò sulle difficoltà, anche

tecnologiche, da superare per realizzare una macchina in grado di pensare oppure, in via subordinata, di ingannare l'osservatore inducendolo a credere di farlo. Nel secondo caso, invece, possiamo solo ipotizzare che Turing si pose i problemi etici che sorgono in seguito alla realizzazione di una macchina pensante. Assumendo infatti di essere in grado di costruirla, saremmo anche disposti ad autorizzarla a pensare? Non è una decisione da poco, perché un'entità pensante è anche senziente e dotata di una qualità formidabile: il libero arbitrio. Che senso avrebbe, altrimenti, costruire una macchina che pensa, per poi perimetrarne il contesto dei pensieri? Le prime domande che questa macchina si porrebbe sarebbero sulla sua identità e sulle sue finalità: chi sono? chi e perché mi ha creato? cosa devo o posso fare?

Storicamente, un'analisi sulle macchine pensanti fu fatta un anno prima da parte di Geoffrey Jefferson, che nel suo articolo "La mente di un uomo meccanico" si pose l'obiettivo di analizzare e comprendere le relazioni tra mente e cervello, partendo dalla grande somiglianza tra le macchine elettroniche e il sistema nervoso. Jefferson riprese ciò che Cartesio aveva affermato duecento anni prima paragonando il comportamento degli animali a quello degli umani. Curiosamente, Cartesio portò ad esempio il pappagallo, in grado di ripetere solo ciò che gli è stato insegnato senza essere in grado di esprimere i suoi propri pensieri. Recentemente, ChatGPT è stato paragonato proprio a un pappagallo "stocastico", capace cioè di esprimersi con sequenze di parole calcolate statisticamente, spesso corrette dal punto di vista semantico, a volte totalmente errate al punto da essere chiamate allucinazioni. Jefferson anticipò e ispirò Turing con argomenti che appaiono attuali quando, ad esempio, afferma che le macchine potranno risolvere solo problemi per i quali sono state istruite,

VINCENZO AMBRIOLA



Avvenire

senza andare oltre. La definizione corrente di intelligenza biologica, e anche quella di intelligenza artificiale, richiede la capacità di risolvere problemi mai visti prima o, in caso contrario, visti in contesti diversi.

L'intelligenza artificiale generale è l'estensione dell'intelligenza artificiale a livelli propri degli umani. Non si tratta di un'intelligenza "stretta", capace ad esempio di giocare a scacchi o di riconoscere immagini, ma in grado di applicare schemi di risoluzione di problemi a qualsiasi situazione. Le origini di questa definizione sono dibattute, ma si ritrovano in un articolo del 1964 di Irving John Good, un matematico britannico che aveva lavorato con Turing a Bletchley Park, perseguito quando, al termine della Seconda guerra mondiale si erano entrambi trasferiti all'Università di Manchester. Good speculò sulle caratteristiche della prima macchina ultra-intelligente, in grado di «superare di gran lunga tutte le attività intellettuali di qualsiasi essere umano, dotato di qualsiasi livello di intelligenza». Si tratta di una macchina che, a buon titolo, può essere considerata pensante e intellettualmente superiore agli umani.

Good non si pose il problema etico della sua realizzazione, la seconda interpretazione della domanda di Turing, ma quello tecnologico e soprattutto matematico. Questo domanda se la pose, invece, Ray Kurzweil quando analizzò ciò che una macchina ultra-intelligente, dotata di intelligenza artificiale generale, potrebbe fare. Lo scenario apocalittico previsto da Kurzweil vede questa macchina studiare il progetto con cui è stata realizzata per migliorarlo e produrre in un tempo finito, una macchina più potente. A sua volta, questa macchina potrebbe ripetere lo stesso processo ma in un intervallo temporale più breve. L'iterazione di questi processi migliorativi farebbe nascere una macchina così potente da prendere il controllo del mondo e (forse) distruggere la specie umana. Kurzweil definì "singolarità" il preciso momento in cui avrà preso forma la prima macchina ultra-intelligente e, quindi, pensante. Veniamo ai giorni nostri e osserviamo ciò che sta accadendo. Un anno fa OpenAI ha messo a disposizione di tutti noi ChatGPT 3.5, un'intelligenza artificiale generativa in grado di dialogare come un essere umano. Molti affermano che questa macchina ha superato di larga misura il test di Turing. Se avete qualche dubbio, provate a interagire con lei e capirete immediatamente il perché di questa affermazione. Ma ChatGPT non è il vero obiettivo che Sam Altman, il fondatore di OpenAI, aveva e ha ancora in mente. Altman crede fermamente nella possibilità di ottenere l'intelligenza artificiale generale e continuerà a operare in questa direzione. Per il momento, però, anziché perseguire l'intelligenza artificiale generale Altman sembra essere impegnato a gestire un'incredibile fonte di profitto, supportato e finanziato dalla Microsoft. Ma poiché i sogni sono talmente da raggiungere, non si può dubitare del suo impegno tecnologico e visionario. Per concludere, dobbiamo credere che un giorno le macchine saranno in grado di pensare? Difficile rispondere, perché il progresso tecnologico punta sempre verso obiettivi lontani e apparentemente irraggiungibili. Inoltre, non sarà scontato capire se e quando questo obiettivo sarà raggiunto, perché tutto dipende da ciò che siamo disposti ad accettare come prova effettiva di macchina pensante. Ad esempio, il test di Turing non ci può più aiutare in questo. Andiamo oltre. Dopo averle costruite, queste macchine saranno anche autorizzate a pensare liberamente? Molto probabilmente questo è il vero argomento etico da affrontare fin da ora. Se ne vedono già alcune tracce nel regolamento

Avvenire

sull'intelligenza artificiale che l'Europa si accinge ad approvare. Il concetto di rischio prevede che alcune pratiche di intelligenza artificiale siano rigorosamente proibite e che quelle ad alto rischio siano tenute sotto stretto controllo. Non è difficile prevedere che, quando le prime scintille di intelligenza generale cominceranno a sprizzare, saranno numerosi i pompieri che accorreranno a spegnerle. RIPRODUZIONE RISERVATA Il progetto "Bloodline Shrine" (2018) di Meredith Monk, curato da Caterina Molteni, presentato al Pio Istituto delle Sordomute di Bologna per la 12ª Art City.