



Filosofia del linguaggio 2023-2024 n° 3

Prof. Stefano Gensini (email: stefano.gensini@uniroma1.it)

Macchine, mente, linguaggio

- Possono le macchine pensare?
- La domanda è posta nel 1950 dal matematico inglese Alan Turing (1912-1954), il quale elabora un test divenuto famoso: se una macchina è in grado di superarlo positivamente, la macchina «può pensare»



A sx Turing; a dx, J. von Neumann, cui si deve la definizione dell'architettura logica dei primi computer

- Test di Turing (dalla Treccani online)
- Si tratta di una situazione sperimentale consistente in un gioco a tre partecipanti: un uomo *A*, una donna *B* e una terza persona *C*. Quest'ultimo è tenuto separato dagli altri due e, tramite una serie di domande, deve stabilire qual è l'uomo e quale la donna. Dal canto loro, anche *A* e *B* hanno dei compiti: *A* deve ingannare *C* e portarlo a fare un'identificazione errata, mentre *B* deve aiutarlo. Affinché *C* non possa disporre di alcun indizio (come l'analisi della grafia o della voce), le risposte alle domande di *C* devono essere dattiloscritte o similmente trasmesse. Il test di Turing si basa sul presupposto che una macchina si sostituisca ad *A*. In tal caso, se *C* non si accorgesse di nulla, la macchina dovrebbe essere considerata intelligente, dal momento che sarebbe indistinguibile da un essere umano. La macchina cioè dovrebbe essere considerata come dotata di una "intelligenza" pari a quella dell'uomo.

La stanza cinese di John Searle

- ▶ Nel 1980 il filosofo statunitense Searle (1932-), autorevole studioso della mente e del linguaggio, ripropone il test di Turing alla luce della funzione acquisita dai computers nella vita moderna.
- ▶ Il 1980 vede (fra l'altro) la nascita dei primi PC, funzionanti in base al S.O. «DOS», che solo dal 1985 in poi comincerà a essere sostituito da Windows.



Cfr. *Menti, cervelli, programmi*,
a c. di G. Tonfoni, CLUP_CLUED

Searle distingue una Intelligenza artificiale (IA) «debole» e una «forte»: la sua critica è rivolta verso la 2nda. Il suo obiettivo è dimostrare che una IA non può in alcun modo *pensare* nel senso comunemente annesso a tale parola. A tal fine immagina la seguente situazione sperimentale (da *Le Scienze*, 259, 1990):

sona. Eccola. Si consideri una lingua che l'individuo in questione non conosce. Io, per esempio, non conosco il cinese: ai miei occhi la scrittura cinese si presenta come una serie di scarabocchi senza significato. Supponiamo ora che io mi trovi in una stanza contenente scatole piene di ideogrammi cinesi e supponiamo che mi venga fornito un manuale di regole (scritto nella mia lingua) in base alle quali associare ideogrammi cinesi ad altri ideogrammi cinesi. Le regole specificano senza ambiguità gli ideogrammi in base alla loro forma e non richiedono che io li capisca. Le regole potrebbero essere di questo tipo: «Prendi uno scarabocchio dalla prima scatola e mettilo accanto a uno scarabocchio preso dalla seconda scatola».

Supponiamo che fuori dalla stanza vi siano delle persone che capiscono il cinese e che introducano gruppetti di ideogrammi e che, in risposta, io manipoli questi ideogrammi secondo le regole del manuale e restituisca loro altri gruppetti di ideogrammi. Ora il manuale con le regole è il «programma di calcolatore», le persone che l'hanno scritto sono i «programmatori» e io sono il «calcolatore». Le scatole piene di ideogrammi sono la «base di dati», i gruppetti di ideogrammi che mi vengono forniti sono le «domande» e quelli che io restituisco sono le «risposte».

Supponiamo ora che le regole del manuale siano scritte in modo tale che le mie «risposte» alle «domande» non si possano distinguere da quelle di una persona di lingua madre cinese. Per esempio, gli individui situati al di fuori della stanza mi potrebbero passare degli ideogrammi il cui significato, a me sconosciuto, sia: «Qual è il colore che preferisci?» e, seguendo le regole, io potrei restituire loro degli ideogrammi il cui significato, a me del pari sconosciuto, sia: «Il colore che preferisco è l'azzurro, ma mi piace molto anche il verde.» Io supero così il test di Turing per la comprensione del cinese, eppure ignoro completamente questa lingua. E, nel sistema che ho descritto, non potrei in nessun modo giungere a capire il cinese, perché non avrei la possibilità di apprendere il significato di alcun simbolo. Come un calcolatore, io manipolo simboli, ma non annetto a questi simboli alcun significato.

Questo esperimento concettuale dimostra che se io non capisco il cinese per il solo fatto di eseguire un programma per la comprensione del cinese, allora non ci riesce alcun altro calcolatore digitale che si limiti a far girare un programma del genere. I calcolatori digitali si limitano a manipolare simboli formali secondo le regole contenute nel programma.



La persona nella stanza non conosce il cinese e si limita a associare serie di ideogrammi in base alle istruzioni date dal manuale.

Ciò che vale per il cinese vale anche per altre attività cognitive. La sola manipolazione dei simboli non basta di per sé a garantire l'intelligenza, la percezione, la comprensione, il pensiero e così via. E poiché i calcolatori sono per loro

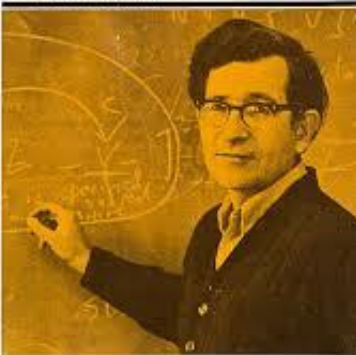
Cognitivismo di
prima
generazione:
mente =
software;
cervello =
hardware

APO 23
Modern Studies in Philosophy

54.05

ON NOAM CHOMSKY: CRITICAL ESSAYS

Edited by Gilbert Harman



A Doubleday Anchor Original

natura dispositivi per operare sui simboli, la semplice operazione di far girare il programma non è garanzia sufficiente di attività cognitive.

Questo semplice argomento confuta in modo radicale le pretese dell'IA forte. La prima premessa dell'argomentazione asserisce semplicemente il carattere formale di un programma di calcolatore. I programmi sono definiti in termini di manipolazioni di simboli e i simboli sono enti puramente formali, cioè «sintattici». È il carattere formale del programma, per inciso, che rende così potenti i calcolatori: lo stesso programma può essere eseguito su una varietà illimitata di calcolatori e un certo complesso circuitale può far girare una gamma illimitata di programmi di calcolatore. Riassumerò questo «assioma» nel modo seguente:

Assioma 1. I programmi di calcolatore sono formali (sintattici).

Questo punto è talmente importante che merita una spiegazione più particolareggiata. Un calcolatore digitale elabora informazione codificandola in primo luogo nel simbolismo che esso usa e poi manipolando i simboli secondo un insieme di regole enunciate con precisione. Queste regole costituiscono il programma. Per esempio nella prima teoria dei calcolatori, enunciata da Turing, i simboli erano semplicemente 0 e 1 e le regole del programma erano del tipo: «Scrivi uno 0 sul nastro, spostalo a sinistra di una casella e cancella un 1.» La cosa straordinaria è che qualsiasi informazione che possa essere espressa in una lingua può essere anche codificata in un sistema del genere e che qualsiasi compito di elaborazione dell'informazione che possa essere risolto con regole esplicite può essere programmato.

gramma ha una sintassi, ma non ha una semantica.

L'assioma successivo serve solo a ricordarci il fatto ovvio che gli atti di pensiero, di percezione, di comprensione e così via hanno un contenuto mentale. Grazie al loro contenuto possono concernere oggetti e situazioni del mondo esterno. Se il contenuto coinvolge una lingua, accanto alla semantica ci sarà una sintassi, ma la comprensione linguistica richiede almeno un ambiente semantico. Se, per esempio, penso alle ultime elezioni presidenziali, la mia mente sarà attraversata da certe parole, ma esse riguardano le elezioni solo perché a queste parole annetto significati particolari in conformità con la mia conoscenza della lingua in cui mi esprimo. Sotto questo profilo esse sono per me ben diverse dagli ideogrammi cinesi. Riassumerò così questo assioma:

Assioma 2. La mente umana ha contenuti mentali (una semantica).

Voglio ora aggiungere quanto è stato dimostrato dall'argomento della stanza cinese. Il possesso dei soli simboli, della sola sintassi, non è sufficiente per possedere la semantica. Le semplici manipolazioni dei simboli non bastano per garantire la conoscenza del loro significato. Riassumerò tutto ciò nel seguente assioma:

Assioma 3. La sintassi di per sé non è condizione essenziale, né sufficiente, per la determinazione della semantica.

A un certo livello questo principio è vero per definizione. Naturalmente sarebbe possibile definire in modo diverso i termini sintassi e semantica. Il fatto è che c'è una differenza fra gli elementi formali, che non hanno un significato o un contenuto intrinseco, e i fe-

nomeni che hanno contenuto intrinseco. Da queste premesse segue la

Conclusione 1. I programmi non sono condizione essenziale né sufficiente perché sia data una mente.

Questo è solo un modo diverso per affermare che l'IA forte è falsa.

È importante capire che cosa esattamente dimostra e che cosa non dimostra questo ragionamento.

In primo luogo, non ho cercato di dimostrare che «un calcolatore non può pensare». Poiché tutto ciò che può essere simulato per via computazionale può essere descritto come un calcolatore e poiché il nostro cervello può a certi livelli essere simulato, ne segue ovviamente che il nostro cervello è un calcolatore ed esso può certo pensare. Ma dal fatto che un sistema possa essere simulato per mezzo della manipolazione di simboli e dal fatto che esso sia pensante non ne segue che pensare equivalga all'esecuzione di manipolazioni formali su simboli.

In secondo luogo, non ho cercato di dimostrare che solamente i sistemi a base biologica, come il nostro cervello, possono pensare. Per il momento questi sono gli unici sistemi conosciuti che di fatto pensino, ma potremmo scoprirne altri nell'universo capaci di produrre pensieri coscienti e potremmo addirittura riuscire a costruire in futuro sistemi artificiali in grado di pensare. Ritengo che questo problema sia aperto a ogni soluzione.

In terzo luogo, la tesi dell'IA forte non sostiene che, per quanto si sappia, calcolatori con programmi giusti potrebbero pensare e avere proprietà psicologiche finora non osservate; la tesi sostiene che essi vanno considerati pensanti

La frontiera mobile della IA



Possono i computer riconoscere emozioni? Possono simularle? Possono avvertirle? Tra tecnologia e fantasia.

Nel corso della storia, le dottrine scientifiche che in Occidente hanno trattato gli uomini semplicemente come parte dell'ordine fisico e biologico comune sono state spesso contrastate da manovre di retroguardia d'ogni sorta. Copernico e Galileo furono combattuti perché sostenevano che la Terra non è collocata al centro dell'universo; Darwin fu combattuto perché affermava che l'uomo discende da animali inferiori. Il modo migliore per inquadrare l'IA forte è di vederla come uno degli ultimi sussulti di questa tradizione antiscientifica, perché nega che esista nella mente umana qualcosa di sostanzialmente fisico e biologico. Secondo l'IA forte, la mente è indipendente dal cervello: è un programma di calcolatore e come tale non è legata ad alcun specifico substrato circuitale.

Molti di coloro che nutrono dubbi sulla portata psicologica dell'intelligenza



Molti di coloro che nutrono dubbi sulla portata psicologica dell'intelligenza artificiale pensano che i calcolatori potrebbero sì capire il cinese e ragionare sui numeri, ma non potrebbero fare certe cose squisitamente umane, cioè (e qui segue la specialità umana che preferiscono): innamorarsi, avere il senso dell'umorismo, percepire l'angoscia della società postindustriale nell'era del tardo capitalismo, o quant'altro. Ma i ricercatori di IA obiettano, giustamente, che così si sposta via via il traguardo. Appena all'intelligenza artificiale riesce una simulazione, essa cessa di avere importanza psicologica. In questo dibattito nessuno dei due contendenti percepisce la distinzione tra simulazione e riproduzione. Per quanto riguarda la simulazione, è facilissimo programmare un calcolatore in modo che scriva «Susi, ti amo»; «Ha ha»; oppure «Soffro l'angoscia della società postindustriale nell'era del tardo capitalismo». Ma è importante rendersi conto che la simulazione non coincide con la riproduzione e l'importanza di questo fatto è la stessa tanto per il pensare di aritmetica quanto per sentire l'angoscia. Non è che il calcolatore arrivi solo fino alla metà campo invece di arrivare fino all'area di rigore. Il calcolatore non parte neppure: non gioca a questo gioco.

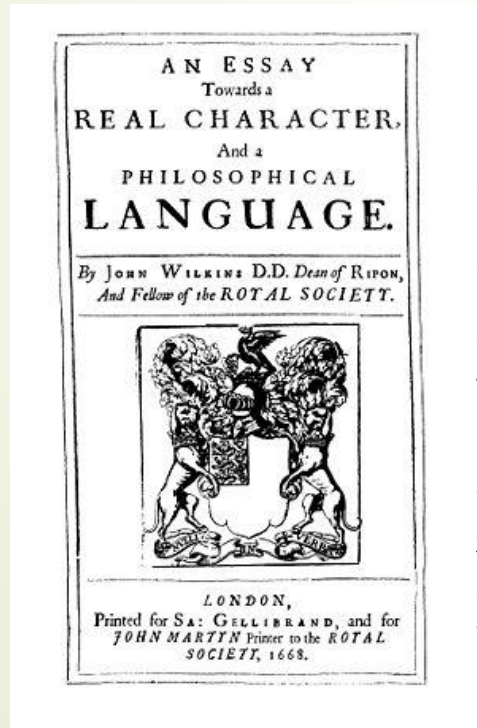


Per concludere (provvisoriamente)...

- ▶ Perché se inserisci il tuo codice fiscale sbagliando una cifra il sistema ti rimanda un messaggio di errore?
- ▶ Come sapete, ci sono programmi di riconoscimento vocale che vi chiedono di rispondere a semplici domande: perché spesso hanno difficoltà a svolgere il loro compito? (Non capiscono ciò che hai detto)
- ▶ Perché, nel riempire al computer un modulo preimpostato il sistema (supponendo che tu abbia un nome troppo lungo) va si blocca se esci dalle apposite caselle?
- ▶ Sai come viene letto dal computer una lettera o un numero che hai appena digitato sulla tastiera?

Una classica sfida per l'intelligenza artificiale: la traduzione automatica

- Precedenti nelle 'lingue universali' del XVII secolo.



Le lingue universali di Joachim Becher e John Wilkins sono liste di corrispondenze fra termini di lingue diverse, equiparate a simboli appositamente ideati.

- Negli anni 1950, anche a seguito delle ricerche sulla crittografia dei tempi di guerra, emerge l'idea di macchine in grado di tradurre automaticamente da una lingua all'altra.
- Warren Weaver (1949), «Translation» su base lessicale
- Nel 1954 a New York, presso la sede di [IBM](#), ebbe luogo la prima dimostrazione pubblica del funzionamento di un [sistema](#) di traduzione automatica in collaborazione con l'[università di Georgetown](#). Durante l'esperimento vennero tradotte 49 frasi dal russo all'inglese con un calcolatore che disponeva di 250 parole di vocabolario e di sei regole grammaticali.



Sviluppi recenti determinati dalla disponibilità di enormi corpora

- ▶ *Google Translator (2006-)*
- ▶ *DeepL Translator (2017-)*
- ▶ Sono molto migliorati nel corso del tempo, riducendo sempre più la necessità di un massiccio intervento redazionale per la correzione degli errori di comprensione e la messa a punto dei testi prodotti dal pdvista della accettabilità linguistica.
- ▶ Provate anche Voi:
- ▶ Inserite un modo di dire o un proverbio (in italiano o altra lingua) e esaminate la risposta: non mancano le sorprese.
- ▶ Ad es. «A ogni morte di papa» (*Once in a blue moon*) viene tradotto da DeepL «At every death of the pope», va cioè perduto il senso traslato del testo ital.

La nuova frontiera: ChatGPT

- ▶ A sviluppare ChatGPT è OpenAI, la società di AlphaFold e GPT-3 che vanta anche il supporto finanziario di Microsoft.
 - ▶ OpenAI ha sviluppato un modello linguistico di grandi dimensioni (LLM) progettato per un'interazione colloquiale e conversazionale con gli utenti finali.
 - ▶ L'idea base è che ciò rappresenti "un investimento pluriennale e multimiliardario per accelerare le scoperte dell'Intelligenza artificiale e garantire che questi vantaggi siano ampiamente condivisi con il mondo".
- ▶ **Da Network Digital 360 (23.10.23)**
 - ▶ *Cosa è possibile fare con ChatGPT:*
 - *effettuare ricerche*
 - *scrivere codice informatico*
 - *automatizzazione di parti del processo di vendita*
 - *offrire servizi post-assistenza quando i clienti acquistano prodotti*
 - *personalizzare istruzioni*
 - *idee di brainstorming*
 - *razionalizzare e migliorare i processi grazie all'automazione*
 - *tradurre testo da una lingua all'altra*
 - *semplificare il processo di onboarding dei clienti*
 - *Aumentare l'engagement dei clienti, migliorando la fidelizzazione*



Ma Noam Chomsky ci invita a non cadere nella trappola

«Come avrebbe potuto osservare Borges, è comico e tragico allo stesso tempo che tanto denaro e tanta attenzione siano focalizzati su una cosa così irrilevante se confrontata con la mente umana. La nostra mente non è, come ChatGPT e simili, un goffo motore statistico per la corrispondenza di schemi, che ingloba centinaia di terabyte di dati ed estrapola la risposta conversazionale più probabile o la soluzione più probabile a una domanda scientifica. Al contrario, la mente umana è un sistema sorprendentemente efficiente e elegante che opera con piccole quantità di informazioni. La mente umana non inferisce correlazioni brute tra dati, ma prova a fornire spiegazioni. Come avrebbe detto Wilhelm von Humboldt, tramite il linguaggio la nostra mente può fare “un uso infinito di mezzi finiti”, creando idee e teorie di portata universale». (*New York Times*, 8.3.23)



In sostanza neppure ChatGPT va oltre il livello *formale* degli enunciati.

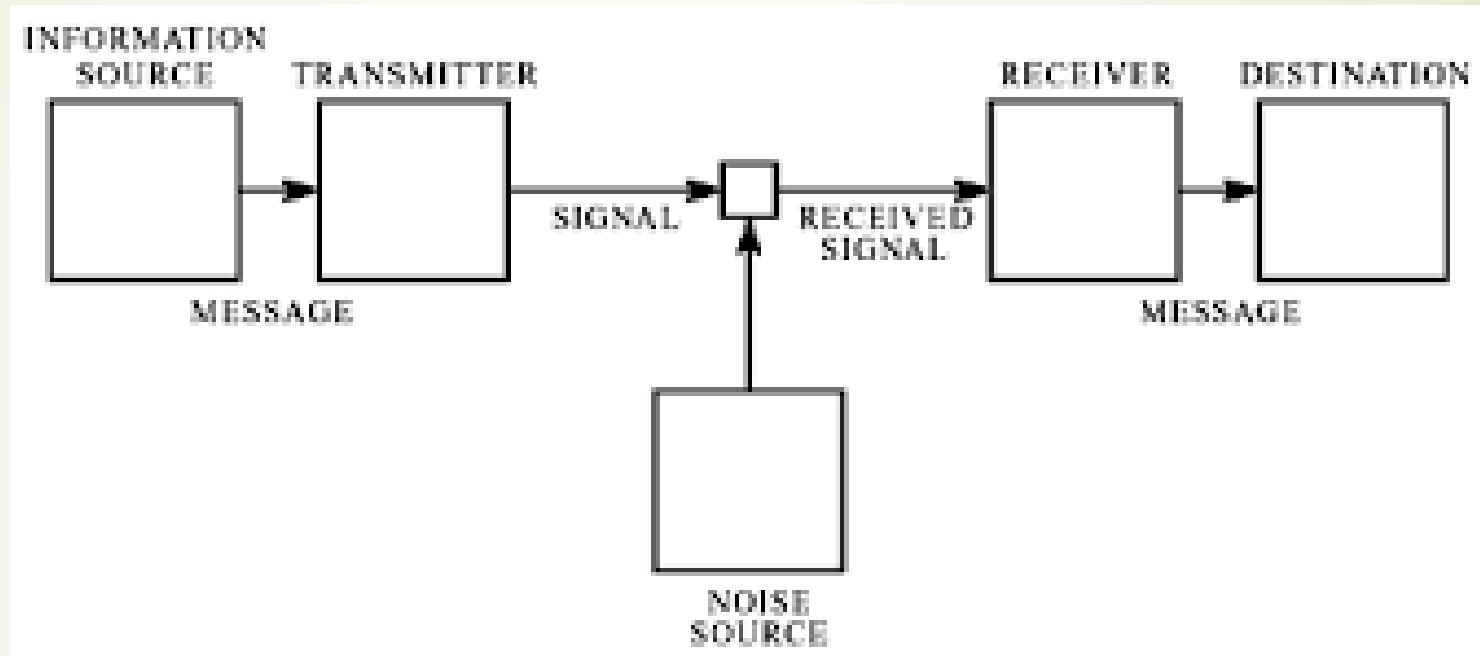
«ChatGPT e programmi simili sono, per loro stessa progettazione, illimitati in ciò che possono “imparare” (vale a dire, memorizzare); sono incapaci di distinguere il possibile dall’impossibile. A differenza degli esseri umani che sono dotati di una grammatica universale che limita le lingue che possiamo imparare a quelle con un certo tipo di eleganza quasi matematica, questi programmi imparano lingue umanamente possibili e umanamente impossibili con uguale facilità. Mentre gli esseri umani sono limitati nei tipi di spiegazioni che possiamo razionalmente congetturare, i sistemi di apprendimento automatico possono apprendere sia che la terra è piatta, sia che la terra è rotonda. Sono fondati semplicemente su probabilità che cambiano nel tempo. (Chomsky, dal *New York Times*, 8.3.2023)



Per concludere, introduciamo il classico modello matematico della comunicazione

- ▶ Esso fu esposto per la prima volta dal matematico Claude Shannon nel 1948, ma ha avuto enorme fortuna, influenzando profondamente anche i successivi modelli.
- ▶ Pensato per la comunicazione fra macchine, esso legge i segni in meri termini simbolici, vale a dire come stringhe di elementi «vuoti», il cui valore ha a che fare con la realtà solo se qualcuno la «traduce» in un linguaggio naturale.
- ▶ Tale modello è tutto fondato su base statistico-probabilistica. Qualsiasi disturbo nella comunicazione (ad esempio del «rumore» che oscura la trasmissione fisica) rende illeggibile/ncomprensibile il messaggio.

Il modello della comunicazione di C. Shannon e W. Weaver



Studiato per la comunicazione telefonica, il modello illustra il funzionamento di una macchina comunicante, ad es. un sistema di allarme. Il messaggio viene codificato in un segnale dal trasmettitore e recepito dal ricevente, ma deve essere decodificato per poter essere utilizzato dall'utente finale (umano)