

Nell'algoritmo tutta sintassi, ma niente significato

Paragonare la mente a un computer è tutt'oggi paradossale: da Castelvechi, i saggi del filosofo del linguaggio John Searle in *Intelligenza artificiale e pensiero umano*

di MARCO MAZZEO

Immaginiamo di essere chiusi in una stanza mentre sfogliamo un testo scritto in cinese: abbiamo a disposizione un manuale che consente di mettere in relazione ogni ideogramma con una serie astratta di simboli convenzionali. Immaginiamo, infine, che l'applicazione ricorsiva di questo meccanismo di traduzione produca un testo finale in perfetto italiano, del tutto fedele all'originale. In un caso del genere, potrei dire di conoscere il mandarino?

Queste poche righe condensano la versione ridotta di uno degli esperimenti mentali più noti del ventesimo secolo, chiamato per l'appunto «la stanza cinese». Del suo autore, il filosofo del linguaggio John Searle, è ora disponibile un volume, *Intelligenza artificiale e pensiero umano* (Filosofia per un tempo nuovo) (traduzione e cura di Angela Condello, Castelvechi, pp. 205, € 20,00) che raccoglie alcuni dei suoi saggi più recenti (scritti tra il 1996 e il 2006) con l'aggiunta di un'intervista inedita.

Searle è una figura unica nel panorama novecentesco: si formò a Oxford, patria di quel ramo della filosofia chiamato «del linguaggio ordinario», ma i suoi saggi non si limitano a descrivere e analizzare il senso comune dei parlanti. Si è confrontato con le ricerche sulla coscienza e sull'intenzionalità di marca fenomenologica; ha offerto una teoria delle istituzioni strettamente legata alla sua visione del linguaggio e della mente, conosciuta come «ontologia sociale»; si è reso protagonista di un serrato dialogo critico con il mondo delle scienze cognitive e, per l'appunto, dell'intelligenza artificiale.

Macchine sprovviste di senso

L'esperimento mentale della «stanza cinese», presente in qualsiasi manuale o storia della ricerca cognitiva, è stato particolarmente discusso, perché è «mentale» (cioè ipotetico o addirittura fantastico) solo in parte. Circa l'eventualità che un computer, opportunamente programmato, realizzi una vera e propria «mente», il filosofo americano dichiara il suo radicale scetticismo: l'operatore presente nella stanza non conoscerebbe affatto il cinese, sarebbe solo l'ingranaggio di un'enorme macchina. Una macchina appariscente, certo, ma non per questo in grado di comprendere il senso del proprio comportamento.

Al contrario, secondo la posizione teorica chiamata «intelligenza artificiale forte», ad avere importanza non sarebbe l'*hardware* di un sistema cognitivo ma solo la struttura del *software* su cui è implementato. Se, ad esempio, un programma di videoscrittura funziona su computer fisicamente diversissimi tra loro (per memoria, modello, velocità di calcolo, marca), perché non pensare a un rapporto simile tra il *software* umano (pensiero, coscienza, mente, linguaggio) e quell'*hardware*



fisico che è il nostro cervello? In fondo, le parole che usiamo e i pensieri che affollano la nostra testa sono gli stessi anche se «girano» su macchine cerebrali differenti. La frase «Mirko e Marco mangiano la mela», ad esempio, mantiene il suo senso tanto che sia io a dirla (o a pensarla) o che a pronunciarla siano Mirko o altri da lui. Questa posizione filosofica è oggi alla ribalta delle cronache perché porta anche a sostenere che se i sistemi cognitivi umani funzionano come un elaboratore di informazioni, in condizioni adeguate un elaboratore di informazioni può eguagliare i sistemi cognitivi dei *sapiens*.

Searle mette in discussione entrambi le affermazioni, proponendo una serie ordinata, comprensibile anche al lettore meno esperto, di aggiornamenti argomentativi circa un esperimento mentale proposto nel lontano 1980. La difesa – ma è in realtà un contrattacco, si muove grosso modo lungo due direzioni. Da un lato, offre un chiaro compendio critico alle obiezioni emerse in questi quarant'anni; dall'altro, risponde in modo ampio a un interrogativo tutt'altro che secondario. Data la velocità dei progressi dell'informatica, siamo proprio sicuri che lo scetticismo circa l'intelligenza artificiale forte sia ancora sostenibile?

La risposta del filosofo è decisamente affermativa. Occorre dichiararsi scettici, perché il miglioramento delle capacità computazionali

non sembra cambiare la questione. Paragonare la mente umana a un computer continua, anche oggi, ad essere paradossale: quest'ultimo è il frutto di una costruzione culturale, dunque *mentale*.

Sarebbe come dire che *I girasoli* di Van Gogh sono tali perché perfettamente corrispondenti alla copia del quadro che si trova nel catalogo di una mostra museale. A tal proposito, Searle discute del rivale dell'ormai celebre sistema di intelligenza artificiale *ChatGPT*, cioè del programma prodotto da Google *LaMDA*, per ribadire la tesi secondo cui «la simulazione non è duplicazione».

La velocità non risolve il problema

La maggiore velocità in megahertz della CPU non sposta il dato di fondo, vale a dire che gli algoritmi di calcolo informatici hanno una *sintassi* ma non una *semantica*. Possono, cioè, organizzare relazioni logiche complesse tra input e operazioni matematiche ma non hanno alcun significato che non sia quello attribuito loro dal programmatore e dall'utente.

Searle non esclude che un giorno sia possibile produrre una vera e propria mente artificiale. Per farlo però, almeno secondo quanto oggi dicono le scienze empiriche, servirebbe produrre un corpo più che una macchina; un organismo parlante più che una struttura in silicio traboccante di velocissimi algoritmi.

La tesi di fondo è che simulare, per esempio, una mente, non equivale a riprodurla