

La stanza cinese: Searle e l'intelligenza artificiale

Nel “ritratto” di John R. Searle mi sono soffermato, dopo una breve presentazione dell’Autore, sul rapporto tra realtà e verità e sulla metafisica di base presentata dal filosofo statunitense nel volume *Mind, Language and Society*. Oggi vorrei soffermarmi sul problema dell’intelligenza artificiale (sull’argomento si veda anche l’intervista al filosofo della scienza Michele Marsonet, sezione *A colloquio con...*).

Secondo Searle, le posizioni su tale questioni sono due: la prima è di coloro che parlano dell’IA come di «uno strumento potentissimo [che] ci permette, ad esempio, di formulare e verificare le ipotesi in un modo preciso e rigoroso» (IA debole); la seconda, invece, è rappresentata da coloro che parlano del «calcolatore» non come di un semplice «strumento per lo studio della mente, ma piuttosto, quando sia programmato opportunamente, è una vera mente» (IA forte). In questo senso – conclude Searle – per l’IA forte «poiché il calcolatore programmato possiede stati cognitivi, i programmi non sono semplici strumenti che ci permettono di verificare le spiegazioni psicologiche: i programmi sono essi stessi delle spiegazioni» (John R. Searle, «Menti, cervelli e programmi», in D. Dennett – D. Hofstadter, *L’io della mente*, Adelphi, Milano, 1985, pag. 341).

Quest’ultima posizione è quella critica dal filosofo statunitense. Per raggiungere il suo scopo Searle supporta la sua argomentazione con un famoso esperimento, l’esperimento della “stanza cinese”. Leggiamo quanto scrive in *Minds, brains and programs*:

«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 (pp. 343-348).

Con questo esperimento Searle vuole dimostrare, dunque, che non è possibile assimilare la mente dell’uomo al computer: la persona nella stanza in realtà *non conosce* il cinese, poiché opera solo sui simboli seguendo delle istruzioni; allo stesso modo, un sistema informatico: la sua attività *riproduce* artificialmente gli atti umani, ma *non possiamo* parlare di atti propriamente umani di comprensione, perché – secondo Searle –, in quest’ultimo caso, entrano in gioco *coscienza* e *intenzionalità* che sono “fenomeni” irriducibili. Infatti, -

scrive il filosofo in *La riscoperta della mente* – «nemmeno appellandoci a una scienza perfetta del cervello saremmo in grado di sottoporre la coscienza a una riduzione ontologica analoga a quella operata dalla scienza contemporanea sul calore, la solidità, i colori o il suono» (p. 131).

Giovanni Covino

bricioledifilosofia.com