

INTELLIGENZA ARTIFICIALE

Uomo e macchina
Parità impossibile?IL «PARADOSSO DELLA STANZA CINESE» PROVEREBBE
CHE MAI UN COMPUTER POTRÀ ESSERE «COSCIENTE»

Alberto Viotto

UNA macchina in grado di pensare è l'obiettivo finale della "Intelligenza Artificiale", una disciplina delle scienze cognitive. A contrastare questo sogno, il filosofo statunitense John Searle ha proposto nel 1980 il "Paradosso della stanza cinese".

Immaginiamo che una persona che non conosce il cinese sia chiusa in una stanza con una serie di regole, scritte nella sua lingua madre, per ordinare in una certa maniera i caratteri della lingua cinese. Queste regole, se seguite scrupolosamente, permettono di rispondere in modo soddisfacente a ogni possibile domanda. Nella stanza vengono introdotti dei fogli con domande scritte in cinese. Utilizzando le istruzioni scritte nella sua lingua, la persona che si trova nella stanza è in grado di compilare in cinese dei fogli con le risposte.

Chi si trova fuori della stanza e vede le risposte correttamente formulate in cinese, immaginerà che nella stanza si trovi una persona che conosce il cinese. Chi è dentro la stanza, però, sa benissimo di non conoscerlo. Secondo Searle, quindi, se anche un giorno esisterà una macchina che ci dia l'impressione di essere in

grado di pensare, intrattenendo con noi una discussione, non si potrà concludere che essa stia effettivamente pensando, perché non farà altro che eseguire una serie di operazioni guidate, esattamente come il finito cinese. A questa macchina mancherà ciò che Searle chiama il "contenuto mentale", un concetto simile a quello di "coscienza".

La "Stanza cinese" ha scatenato una quantità incredibile di commenti e di articoli, e ancora adesso è un argomento molto dibattuto su Internet. Una delle principali osservazioni prende in esame il sistema composto dalla persona nella stanza e dalle istruzioni. A conoscere il cinese non è la persona da sola, ma il sistema composto dalla persona e dalle istruzioni. A questa obiezione Searle ha ribattuto che, se ciò che è dentro la stanza non ha "contenuto mentale", questo non può esserci nemmeno nell'elenco delle istruzioni.

Un'altra obiezione riguarda la velocità con cui la persona dentro la stanza deve eseguire le istruzioni. Per rispondere correttamente alle domande, la quantità di regole da eseguire deve essere enorme. Esistono programmi di Intelligenza Artificiale che sono in grado di dialogare con una persona, sia pure in maniera molto limitata; per farlo, però, devono eseguire milioni di istruzioni elementari al secondo. Come ha scritto Steven Pinker, direttore del centro di scienze cognitive del MIT, "se incontrassimo una persona che sembrasse conversare intelligentemente in cinese, ma in realtà eseguisse in frazioni di secondo milioni di regole memorizzate, negheremmo che capisca il cinese? Non è tanto sicuro."

La differenza tra la velocità

con cui un uomo può eseguire delle operazioni e la velocità richiesta dall'esempio di Searle può cambiare i termini del problema. Secondo Patricia e Paul Churchland, l'argomento di Searle potrebbe essere usato contro la fondamentale teoria di Maxwell delle onde elettromagnetiche, secondo cui la luce consiste di onde elettromagnetiche. Un uomo che tiene in mano una calamita facendola oscillare crea radiazione elettromagnetica, ma non esce luce. Seguendo Searle si concluderebbe che la luce non è radiazione elettromagnetica, ma ciò che fa fallire l'esperimento è l'insufficiente velocità del movimento.

La risposta alla domanda "un calcolatore può pensare?" dipende da ciò che definiamo come "pensiero". Nel 1950 il matematico Alan Turing, uno dei padri della Intelligenza Artificiale, ha proposto un criterio per rispondere a questa domanda. Secondo Turing, per essere chiamata intelligente una macchina deve essere in grado di far credere ad un osservatore esterno di essere una persona. L'interazione tra la macchina e l'osservatore può avvenire tramite una telescrivente o un dispositivo simile, per evitare la necessità di emulare la voce o l'aspetto umano.

Il criterio di Turing deriva da una affermazione di Cartesio, che scrisse nel suo "Discorso sul metodo" (1637): "Si può ben concepire che una macchina sia fatta in modo tale da proferire parole ... ma non si può immaginare che possa combinarle in modi diversi per rispondere al senso di tutto quel

che si dice in sua presenza, come possono fare gli uomini, anche i più ottusi."

Per superare il test di Turing, nell'interazione tra la macchina e l'osservatore non vi possono essere indizi di un comportamento diverso da quello che potrebbe avere una persona; anche investigando sulle emozioni o sui sentimenti di ciò che si ha di fronte non si deve capire che non è un essere umano. La macchina deve emulare completamente il comportamento umano, includendo le emozioni, i sentimenti, la coscienza e tutto ciò che può caratterizzare una persona.

Questo obiettivo è ancora molto lontano, ma se fosse raggiunto potremmo negare che una macchina sia in grado di pensare? Searle pone l'accento sul modo in cui si ottiene un comportamento che si possa ritenere umano, ma non coglie un punto fondamentale: noi non sappiamo che cosa succede nella mente delle altre persone, come non sappiamo che cosa succede dentro la stanza cinese. Per assurdo, non potremmo escludere che tutte le altre persone del mondo siano state sostituite da automi dall'aspetto umano in grado di superare il test di Turing, come in alcuni film di fantascienza. Se una macchina riesce a comportarsi esattamente come un altro uomo, non possiamo fare a meno di affermare che pensi, né più né meno di come affermiamo che gli altri uomini pensano.



Alan Turing