

NONA LEZIONE: coscienza e Intelligenza Artificiale

Se realizzassimo macchine artificiali coscienti probabilmente potremmo capire qualcosa di più della coscienza rispetto ad oggi. Avremmo un modello di confronto e di verifica per la nostra coscienza.

Searle rigetta l'IA nella versione forte che concepisce il cervello come un computer e la coscienza come un programma di computer.

Vi oppone l'IA nella versione «debole» secondo cui il computer è solo un mezzo utile per fare simulazioni della mente, o meglio per simulare i processi cerebrali. Ma la simulazione non va confusa con la realtà effettiva di quei processi. In questo errore, a suo avviso, cade Dennett, che per Searle considera i programmi dei computer implementati come stati mentali, e non come cause che li generano.

Searle sulla coscienza delle macchine

Searle comunque non esclude che un computer possa essere cosciente, nel senso che la coscienza possa essere un processo emergente. Ma la computazione è un processo matematico astratto che esiste solo in virtù di osservatori consci che interpretano la computazione a differenza dei sostenitori dell'IA forte che assumono la computazione come un dato oggettivo.

I computer sono solo calcolatori che si limitano ad eseguire insiemi sintattici di operazioni su simboli formali. Ma la mente non è un dispositivo di calcolo formale, ha contenuti poiché interpreta i simboli formali, ovvero dà loro un significato.

Per Searle sintassi e semantica non vanno confuse come invece, a suo parere, fa Dennett. L'esperimento mentale della "stanza cinese" prova per Searle che possiamo tradurre il testo di una lingua nel testo di un'altra lingua, entrambe sconosciute, se abbiamo una tavola di corrispondenza che per ogni segno della prima lingua ci offre un segno della seconda lingua senza capirne il senso².

Dennett sulla coscienza delle macchine A

Dennett dissente da Searle a proposito della irriducibilità soggettiva della coscienza e perché sostiene l'IA forte. Per Dennett l'IA ci sta aiutando a gettare luce sulla coscienza offrendoci paradigmi e metafore che egli introduce e impiega largamente per superare le impasse del passato e darci un nuovo punto di vista.

In particolare egli si avvale del modello della mente come software rispetto al cervello inteso come hardware. Così egli nega l'esistenza di un correlato neuronale del sé (l'hardware) e afferma che esso esiste solo come entità virtuale, come nostra costruzione culturale che si può far risalire agli eventi narrati nell'Iliade e nell'Odissea. Dennett definisce “la coscienza umana come un sistema di macchine virtuali che si è evoluto geneticamente e memeticamente, per aver ruoli specializzati nella “nicchia cognitiva” costruita dai nostri antenati nel corso di millenni”.

Dennett sulla coscienza delle macchine B

Dennett crede che non ci sia nessuna impossibilità di principio nel realizzare una IA pensante e cosciente. Non lo abbiamo finora fatto poiché non conosciamo esattamente come funziona il cervello. Sappiamo che esso possiede un'architettura massicciamente parallela, ovvero che elabora molto velocemente le informazioni attraverso milioni di canali in parallelo. Non abbiamo ancora la tecnologia per riprodurre tale funzionamento in una macchina. Quanto poi alla coscienza, possiamo presumere di dedicare alcuni canali a registrare e sovrintendere ciò che fanno altri.

La coscienza non contiene elementi irriducibilmente soggettivi, come invece afferma Searle, poiché questo la renderebbe impenetrabile alla scienza. La coscienza è esplorabile in linea di principio poiché in fondo è un sistema computazionale. Perciò la mente sta al programma come il cervello sta all'hardware.

Dennett sulla coscienza delle macchine C

Se avessimo un computer digitale capace di attivare il programma giusto abbastanza velocemente da riprodurre in tempo reale il funzionamento del cervello, realizzeremmo quei poteri causali attraverso cui esso produce il controllo della rapida, intelligente, attività intenzionale. Tuttora non si è ancora riusciti in questo intento e, anzi, nonostante i successi conseguiti si è lontani dall'obiettivo. Il programma giusto è la sintassi attivata, ovvero la semantica, cosicché è superato il dualismo semantica-sintattica di matrice cartesiana in cui cade Searle.

Chalmers sulla coscienza delle macchine

Non lontano da Dennett è Chalmers, che respinge l'argomento della stanza cinese in quanto non rispetta il ruolo centrale dell'implementazione. I programmi sono oggetti computazionali astratti puramente sintattici ma le loro implementazioni sono sistemi concreti con dinamiche causali e quindi non sono puramente sintattici. Un'implementazione ha potere causale nel mondo reale, ed è in virtù di questo potere che la coscienza sorge. Anche se il programma è sintattico la sua implementazione ha contenuto semantico.

Quindi non ci sono impedimenti alle ambizioni dell'IA di realizzare delle macchine coscienti. Diversi tipi di computazione possono bastare per addivenire all'esperienza conscia. La spiegazione computazionale è centrale nello studio della coscienza in quanto può catturare ogni sorta di organizzazione causale, anche quella che causa appunto la coscienza. Rimane aperta solo la questione di quale classe di computazione è sufficiente per riprodurre la mente umana.

Un'applicazione specifica

La filosofa e psicologa evolutiva Allison Gopnik ha proposto un interessante esperimento con l'IA. La Gopnik nel suo *Il bambino filosofo* sostiene che la coscienza dei bambini è radicalmente diversa da quella degli adulti.

Possiamo avvalerci dell'IA per simulare l'apprendimento e il funzionamento della mente del fanciullo e quelli della mente dell'adulto per verificare quali e quante sono le differenze fra le due.

Il ricorso all'IA è consigliato dal fatto che non è facile stabilire come funziona la mente di un fanciullo di 4 anni, e nemmeno si può ricostruire sulla base della testimonianza dell'adulto poiché è avvenuto un cambiamento radicale.

Non escludiamo che le macchine da noi realizzate possano essere autocoscienti

Realizzare una macchina non biologica cosciente non ci sembra assurdo, come evidenzia il grandissimo successo di pubblico e critica del film 2001 *Odissea nello spazio* (1968) di Kubrik e Clarke, che possiamo assumere come un valido esperimento mentale.

L'elaboratore Hall 9000, un protagonista della vicenda, è rappresentato come perfettamente cosciente. Eppure la gente non si è alzata durante la proiezione del film protestando che le veniva proposto un film assurdo.

Anzi il film è stato accolto con entusiasmo dal pubblico ed è ancora considerato un capolavoro. Ora va da sé che è inconcepibile un capolavoro basato su una trama assurda, priva di qualsiasi credibilità. Ogni storia per essere apprezzata universalmente dev'essere giudicata per lo meno verosimile.

IA autocoscienti

I continui progressi dell'IA vanno anche nella direzione dell'auto supervisione e dell'autocontrollo. Nella misura in cui le macchine apprendono e si correggono, si approssimano alla coscienza, che non è qualcosa di divino. Diverse aziende producono umanoidi sempre più capaci di interagire con gli umani, evidenziando non un comportamento uniforme e rigido, ma adattandosi alle loro specifiche esigenze. Così come nel regno animale troviamo diverse forme di coscienza, altrettanto le più raffinate IA si approssimano alla coscienza umana. Auto-ML-Zero è un programma di Google che sviluppa in assoluta autonomia programma di IA.

Diversi programmi di IA ne sviluppano altri sempre più raffinati in piena autonomia. Basta vedere su Youtube, l'intervista al robot Sophia: è sconvolgente.