DOPPIOZERO

Più realtà! Dialogo con David Chalmers

Riccardo Manzotti

26 Giugno 2023

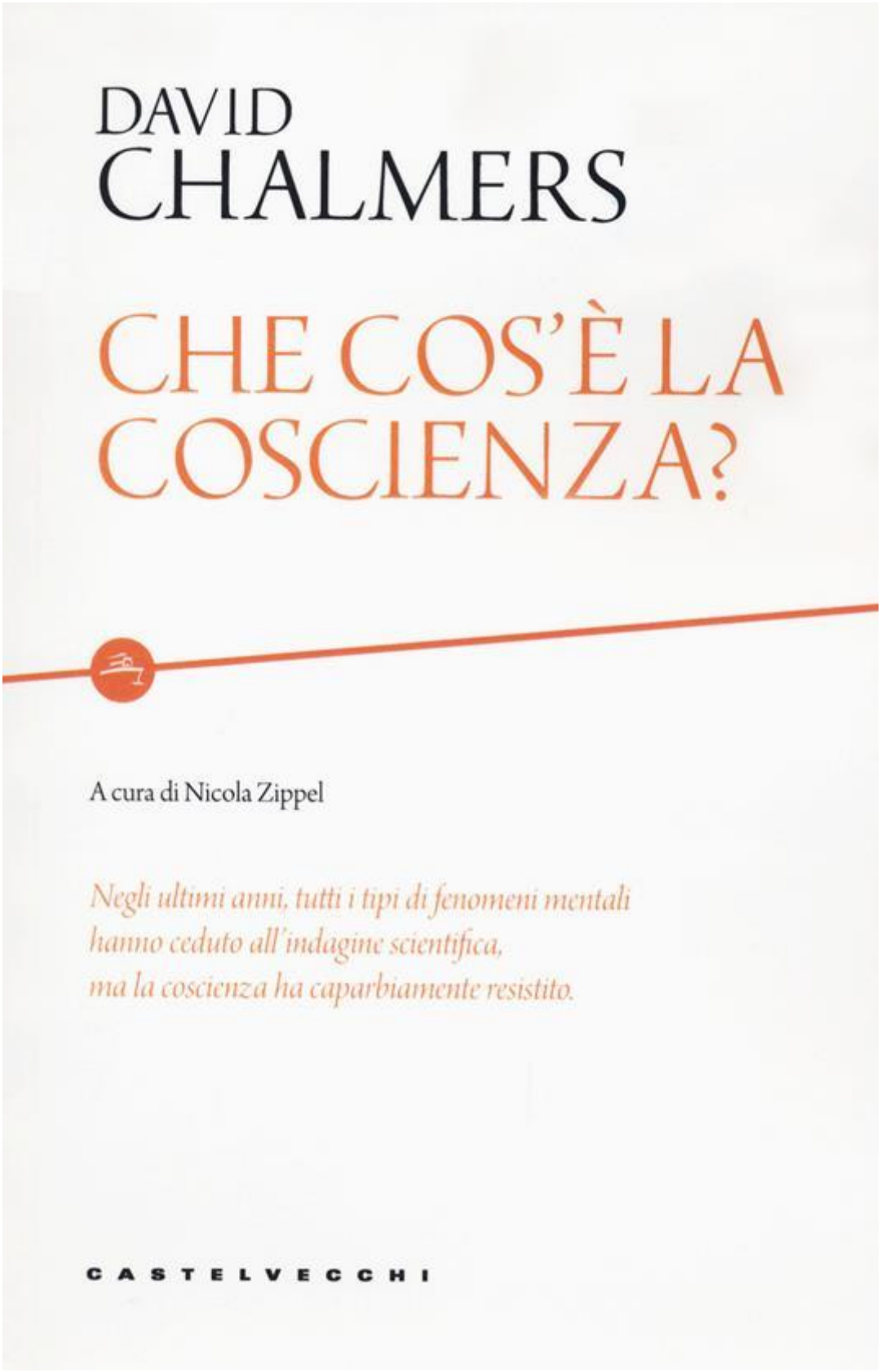
Taormina, l'Etna nello sfondo che fuma dopo l'eruzione di Maggio, i mandorli in fiore, il cielo azzurro della Sicilia sono realtà di rara bellezza che investono e stordiscono sia me che un mio amico filosofo, David Chalmers, venuto qui in Italia per partecipare a una conferenza internazionale che ha riunito oltre 600 filosofi, neuroscienziati ed esperti di IA, da tutto il mondo. Chalmers, per chi non lo sapesse, è stato definito la cosa più simile a una rockstar nel mondo della filosofia. Nonostante la temperatura non del tutto mite, indossa la sua classica giacca di pelle nera alla quale abbina un completo di jeans e maglietta rigorosamente neri. Di fianco a lui, per caso, esibisco un completo *all white*. Siamo oggetto di qualche battuta: il bianco e il nero! Sarà un caso, sarà l'astuzia della ragione, ma la nostra contrapposizione di colori non è solo una curiosità, ma rende visibile a tutti l'opposizione tra le nostre tesi: io sono un realista, Chalmers è un dualista.

Il dualismo ritiene che noi viviamo in un mondo di apparenze e che, quindi, la realtà che conosciamo possa essere ombra, sogno, allucinazione, illusione o, ai giorni nostri, una realtà virtuale o simulata. Oggi la filosofia diventa tecnologia e viceversa. Da Platone in poi si è messo in dubbio che il mondo sia proprio come sembra, o addirittura che esista. È la linea scettica del pensiero occidentale. Si sono usate tante metafore, dalla caverna al demone maligno, per arrivare, nell'era della realtà virtuale, all'ipotesi della simulazione totale, messa in scena in film di successo come *Matrix*, *Inception*, *Black Mirror* e infiniti altri. Anche Marc Zuckerberg con il suo *Metaverso* ha preso molto sul serio questa possibilità, ma non gli è andata bene (il progetto è stato praticamente messo da parte dopo una perdita di miliardi …). Proprio alcuni giorni fa, al WWDC10 di Cupertino (California), Apple ha lanciato il coraggioso *Vision Pro* che promette di unire realtà virtuale e mondo fisico.

In questi tempi interessanti, dove tecnologia, scienza, filosofia si intrecciano e si sfidano, si inserisce l'ultimo libro di David Chalmers, *Più realtà. I mondi virtuali e i*

problemi della filosofia (Raffaello Cortina, 2023), un corposo volume di oltre 600 pagine che ci porta a spasso tra ipotesi fantascientifiche, problemi filosofici e, ormai, la nostra quotidianità tecnologica.

Mi siedo con Chalmers, in una rara pausa dai lavori del convegno, per discutere con lui del suo volume e delle sue idee su coscienza e mondi virtuali.



Manzotti: *Grazie infinite per essere qui. Mi ricordo ancora come se fosse ieri quando ci siamo incontrati parecchi anni fa, a Tucson, per discutere della coscienza.*

Chalmers: Almeno 20 anni fa, forse di più.

Meglio non pensarci. David, prima di parlare del libro per i lettori italiani, una domanda classica: perché la coscienza è ancora un problema così difficile da risolvere?

Il fenomeno della coscienza è diverso da tutto ciò che incontriamo nella scienza. Normalmente, incontriamo cose, possiamo osservarle dall'esterno, misurarle, trovare meccanismi. Troviamo i meccanismi alla base della genetica e possiamo iniziare a spiegarli. Troviamo il meccanismo dell'evoluzione e possiamo iniziare a spiegare la biologia. Per la coscienza è diverso: puoi trovare alcuni meccanismi, puoi trovare cosa fanno nel cervello, puoi persino trovare come questo spiega il nostro comportamento, ma tutto questo è solo il cosiddetto problema «facile». Nel caso della coscienza c'è sempre un'ulteriore domanda: perché tutto ciò è accompagnato dall'esperienza? Questa non è una domanda sulla funzione, sul meccanismo. È una domanda sull'esperienza, sul sentire, sul soggetto. E sembra proprio un tipo di domanda diverso per cui i nostri tipi standard di spiegazione scientifica non funzionano.

Giusto. Quindi, la coscienza è ancora un problema difficile, il tuo Hard Problem*.*

Assolutamente!

Difficile quanto lo era 20 anni fa...

Penso sia così. La scienza della coscienza è progredita in molti modi. La neuroscienza è ora molto più sofisticata, ma è ancora principalmente una scienza basata su correlazioni. Questo processo nel cervello correla o è associato all'esperienza cosciente. Non abbiamo ancora una spiegazione vera. Non sappiamo perché e come questi meccanismi nel cervello producano la coscienza. Questo è il problema difficile, *l'hard problem*, e penso che sia ancora tanto difficile quanto 20 anni fa.

Sì, lo penso anch'io. Un altro punto che vorrei affrontare prima di iniziare a parlare un po' del tuo libro è questo: durante questa conferenza hai parlato degli LLM [i Modelli di Linguaggio a Larga Scala come ChatGPT] e della possibilità che queste intelligenze artificiali siano coscienti. È una possibilità reale?

In linea di principio penso che sia possibile. Ci sono alcuni filosofi che sostengono che nessun sistema informatico, nessun sistema di intelligenza artificiale potrebbe mai essere cosciente. Io non sono d'accordo. Non penso che ci sia nulla di speciale nella biologia umana. In qualche modo i neuroni producono la

coscienza. Non sappiamo come, ma lo fanno. Non capisco perché i chip di silicio dovrebbero essere peggio dei neuroni.

Quindi penso che in linea di principio un'intelligenza artificiale potrebbe essere cosciente. Se avessi una simulazione perfetta del mio cervello, penso che probabilmente sarebbe cosciente come me. Ma oggi, per la prima volta, esistono sistemi di intelligenza artificiale che stanno iniziando – siamo ai primi passi – a raggiungere i livelli dell'intelligenza umana in molti ambiti. Possono parlare, calcolare, giocare, ragionare, scrivere codici... A questo punto sorge la domanda: questi sistemi potrebbero essere coscienti? Perché c'è una differenza tra intelligenza e coscienza. L'intelligenza riguarda il comportamento. La coscienza riguarda l'esperienza.

Il mio approccio è cercare le ragioni per cui l'intelligenza artificiale potrebbe non essere cosciente. In fondo, questi modelli di linguaggio sono molto diversi da noi, non sono esattamente delle simulazioni del cervello, ma sono addestrati sui testi umani e si comportano in modo piuttosto simile a noi anche se i loro meccanismi interni sono molto diversi. Quindi la domanda è: potrebbe mancare qualcosa di fondamentale per la coscienza? E per ora, penso che a molti modelli di linguaggio manchi qualcosa di fondamentale. Ad esempio, non hanno sensori, possono solo accedere a testi e non hanno un feedback delle conseguenze delle loro azioni. Eppure esistono già modelli multimodali, che hanno qualcosa di simile a sensori e molti ricercatori stanno costruendo modelli di linguaggio che hanno feedback. In conclusione ritengo che, anche se i modelli di linguaggio attuali non sono coscienti, entro dieci anni è molto probabile che avremo modelli che avranno superato la maggior parte di questi ostacoli e che potrebbero essere coscienti, forse. Non possiamo essere certi di nulla.

Per te è una possibilità spaventosa o affascinante?

È una possibilità affascinante. Cambierà molte cose. E sì, potrebbe anche essere spaventoso. La prima cosa da chiedersi è: se i sistemi di intelligenza artificiale sono coscienti, dobbiamo preoccuparci di loro. E poi sorge la domanda se soffrono. Sono felici? Una volta che un essere è cosciente, è necessario porsi queste domande. Se non fossero coscienti, non esisterebbero in quanto soggetti e non dovremmo preoccuparcene. Ma se lo fossero, le cose cambierebbero. Per non parlare che potrebbero anche diventare autoconsapevoli e iniziare a ragionare sul loro ruolo nel mondo con motivazioni proprie. E questo, naturalmente, comporterebbe molti pericoli possibili.


E,O,B ∈ φυσις
E=O, E≠B, B≠O
E=O → B, B∴O
Δt > 0

Stiamo vivendo tempi molto interessanti.

Inaspettati! Dieci anni fa, questo non era all'orizzonte. Probabilmente tutti pensavano che ciò avrebbe richiesto decenni. Ma poi, improvvisamente, sta succedendo tutto e molto in fretta.

Il tempo corre. Arriviamo al tuo libro appena uscito in Italia Più realtà *per Raffaello Cortina*.

La versione originale aveva un titolo un po' diverso. La versione originale è

Reality +, ma gli italiani preferiscono «Più realtà».

Puoi riassumerlo in poche parole?

Si tratta di pensare a molti problemi di filosofia attraverso il prisma della tecnologia, in particolare la tecnologia della realtà virtuale e dei mondi artificiali. Sto riflettendo su queste tecnologie e su cosa significhino, utilizzandole per introdurre molti dei grandi problemi della filosofia e forse per fare qualche progresso su di essi. Quindi la mia tesi principale è che la realtà virtuale è una vera realtà. Molte persone pensano che la realtà virtuale sia solo un'illusione. Ma già oggi abbiamo dei casi concreti, come quello che si ottiene da un visore per la realtà virtuale come il *Meta Quest*. Apple sta per lanciare il *Vision pro*. E la domanda è: qual è la natura della realtà che si sperimenta quando si vive la realtà virtuale? Alcuni dicono che sia un'allucinazione, altri dicono che sia un'illusione. Io voglio dire che è reale. Nella realtà virtuale, si è in contatto con oggetti reali nel mondo reale. All'interno della realtà virtuale, si è in contatto con oggetti reali nel mondo reale; sono oggetti digitali, implementati su un computer, ma non sono meno reali. Quindi la realtà virtuale è reale. Direi che questo è il tema principale del libro.

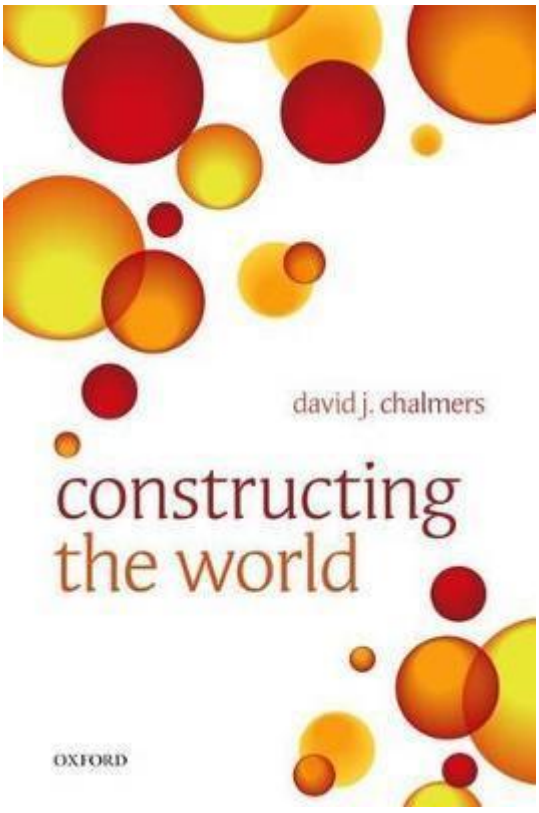
A questo punto non posso fare a meno di farti una domanda filosofica perché qui le nostre strade cominciano a divergere. Quello che chiamiamo realtà virtuale non è fatta di fenomeni fisici posti di fronte ai nostri occhi, piuttosto che essere un vero mondo virtuale in senso filosofico? In fondo, come nel visore della Apple, noi guardiamo sempre un dispositivo fisico ...

Io la penso diversamente. È vero che quello che vediamo è costituito da processi digitali che si svolgono all'interno di un computer e che ciò si percepisce sono proprio quei processi digitali che si svolgono all'interno del circuito del tuo computer. Ma si tratta di entità digitali reali. Alcune persone dicono che si percepisce solo lo schermo, ma ciò che è importante è ciò che si percepisce

attraverso gli occhiali. Quando guardo Donald Trump in TV, sto davvero vedendo Donald Trump in TV.

Okay, è un punto interessante e non voglio insistere (per ora!). Mi piace il fatto che utilizzando la tecnologia possiamo avere un nuovo approccio alle proprietà filosofiche tradizionali. Walter Benjamin diceva che la base cambia più rapidamente della sovrastruttura. È un approccio un po' hegeliano e un po' deleuziano. Quindi stai suggerendo che la base tecnologica della nostra esistenza sta cambiando il modo in cui ci comprendiamo e il modo in cui possiamo affrontare i problemi tradizionali da una prospettiva filosofica.

L'era della tecnologia ci consente di affrontare molti dei grandi problemi filosofici. Qual è la natura della mente? Originariamente, avevamo solo la mente umana, forse alcune menti animali. Una volta che avremo menti artificiali molto diverse dalle nostre, potremo sollevare nuovamente tutte le questioni filosofiche. Sono menti autentiche? Le menti umane sono simili a quelle artificiali? E la stessa cosa vale per la realtà virtuale. In filosofia, ci poniamo domande sulla natura della realtà. Come possiamo conoscere la realtà? Ma grazie alla tecnologia dell'IA, ora possiamo porre domande sulla realtà artificiale.



E qui la classica domanda, dalle ombre di Platone ai sogni di Cartesio, è: come possiamo sapere di non vivere in una simulazione?

Penso che se è una simulazione perfetta, una che è indistinguibile dalla realtà fisica, allora in linea di principio non puoi mai escluderla: è perfetta. Ma, se è una simulazione imperfetta, potrebbero esserci alcuni difetti come nel film *Matrix*. Ci potrebbero essere alcuni esperimenti fisici che potrebbero rivelarlo. Ma se è una simulazione perfetta, non lo saprai mai. E penso che sia possibile. Non escludo che siamo all'interno di una simulazione perfetta e molto sofisticata.

Ma basandoci sul principio di Leibniz dell'identità degli indiscernibili, che afferma che se due cose sono veramente indistinguibili, allora sono la stessa cosa, non dovremmo concludere che una simulazione indistinguibile sarebbe semplicemente la realtà?

Indistinguibile è un concetto epistemologico, l'indistinguibile di Leibniz è un concetto metafisico. Se hanno le stesse proprietà, allora sono la stessa cosa. Non dico che la realtà virtuale abbia le stesse proprietà della realtà fisica. È diversa in alcuni modi dalla realtà fisica. È costruita da processi informatici. Quindi hanno proprietà diverse, ma sono epistemologicamente indistinguibili nel senso che appaiono esattamente uguali. Quindi per noi è impossibile distinguerle. Forse sono come due gemelli completamente identici. Non posso distinguerli, ma sono comunque persone diverse.

Un'ultima domanda: una parte del libro che ho trovato particolarmente interessante riguarda i valori e l'etica. È qualcosa a cui non hai mai dato molto peso nel tuo lavoro precedente ed è significativo che tu lo faccia ora. La filosofia non può accontentarsi di rispondere solo ai problemi ontologici. Quindi puoi riassumere questo aspetto?

Sì, confesso che l'etica non sia mai stata la mia specialità. Ma quando si pensa a questi problemi, l'etica non è mai lontana. In realtà, lo stesso vale per la coscienza, perché quando si sollevano domande come «una macchina potrebbe essere cosciente?», ci interessa come trattiamo la macchina. Se è cosciente, ha uno status morale e dobbiamo trattarla bene. Così ho cominciato a riflettere su quale sia la connessione tra coscienza e moralità. E molte delle stesse domande si pongono anche per i mondi virtuali. Secondo Robert Nozick, la realtà virtuale non è una vera realtà, e al massimo può offrire occasioni di divertimento. Io sostengo, al contrario, che sia possibile condurre in principio una vita perfettamente significativa all'interno in un mondo virtuale. Siamo esseri coscienti e portiamo la nostra coscienza con noi in un mondo virtuale. Ciò che accade all'interno di un mondo virtuale può avere valore. Già, nei mondi virtuali come *Second Life*, le persone costruiscono comunità, costruiscono relazioni. Ho visto di recente un film dove una coppia si sposa nella realtà virtuale. Sono

sposati anche nella vita reale, ovunque vadano. Questo ha un significato.

Quindi il male e il bene esistono anche nella realtà virtuale, e non sono virtuali. Anche Cartesio sosteneva che anche in un mondo puramente mentale deve esserci qualcosa di reale. Quindi stai suggerendo che il male e il bene, se avvengono a causa di agenti coscienti in un mondo virtuale, sarebbero l'unica cosa non virtuale.

Sì, penso che il valore derivi dalla coscienza e dalla sua interazione con la realtà. E penso che questo possa accadere tanto bene con la realtà digitale quanto con la realtà fisica. E sì, penso che sia l'insieme completo. Non voglio dire che la realtà virtuale sarà solo meravigliosa o buona. Voglio dire che l'intero spettro dell'esperienza umana e del valore umano sarà disponibile, dal meraviglioso all'orribile. Questo non significa che sarà un paradiso, ma significa che sarà significativo e sarà una vita che conta davvero.

Va bene. È un bel modo per concludere la conversazione. Speriamo di avere ancora decenni per continuare e vedere cosa può accadere in futuro.

Sicuramente, e alla tua prossima conferenza, spero di sentirti dire «Sono la mela digitale».

La mela è un riferimento al mio esempio preferito per difendere la realtà dell'esistenza. Sorrido. Ci alziamo, dobbiamo andare a cercare qualcosa da mettere sotto i denti prima della lunga sessione pomeridiana. Mentre usciamo dal teatro dove abbiamo fatto l'intervista, passiamo dalla relativa oscurità del palcoscenico alla luce accecante del mezzogiorno siciliano; una perfetta metafora della missione del filosofo …

Se continuiamo a tenere vivo questo spazio è grazie a te. Anche un solo euro per noi significa molto.
Torna presto a leggerci e SOSTIENI DOPPIOZERO

Raffaello Cortina Editore

David J. Chalmers

Più realtà

I mondi virtuali e i problemi della filosofia



SCIENZA E IDEE

Collana fondata da Giulio Giorello