

Beautiful Mind *L'informatico Manuel Blum*

La coscienza dei robot

PIERGIORGIO ODIFREDDI

A cercare di definire la coscienza sono stati dapprima i filosofi, poi gli psicologi e infine i neurofisiologi. Ora, nell'era digitale dei "cervelli elettronici", è naturale che anche gli informatici se ne interessino, e abbiano qualcosa di significativo da dire al riguardo: ad esempio, sul problema se le macchine possano a loro volta essere coscienti. Uno di coloro che ci lavorano è l'americano Manuel Blum, vincitore del premio Turing nel 1995 per la teoria astratta della complessità di calcolo dei programmi, col quale abbiamo parlato al recente Meeting delle medaglie Fields di Heidelberg.

Perché cercare una definizione informatica della coscienza?

«Perché l'informatica ha un vantaggio nei confronti delle discipline umanistiche: permette di dare definizioni precise, sulle quali si può o no convenire, ma dalle quali si possono poi derivare conseguenze verificabili o refutabili in maniera sperimentale».

E cosa sta cercando di fare, precisamente?

«Poiché è più facile capire cose piccole e semplici che cose complicate e grosse, cerco di costruire un modello minimale che mostri qualche livello di coscienza, senza pretendere di coprire la gran varietà di manifestazioni cosce che sono state studiate».

Come è arrivato a interessarsi di questi problemi?

«Quand'ero bambino mio padre mi disse che se volevo dimostrare di essere bravo, avrei dovuto cercare di spiegare come funziona il cervello. Al momento di iscrivermi all'università scelsi ingegneria elettronica, perché pensavo di fare robotica. Ma nel frattempo iniziai anche a seguire

corsi sulle *Opere complete* di Freud, perché credevo che la psicoanalisi fosse la via giusta per capire il cervello».

E cosa accadde?

«Mi accorsi che non era proprio così. E quando il professore di cui seguivo i corsi vide che ero

insoddisfatto, mi indirizzò dal neurofisiologo Warren McCulloch, che era considerato l'anti-Freud. Nel 1943 aveva scritto un saggio su *Un calcolo logico delle idee immanenti nell'attività nervosa*, insieme al matematico Walter Pitts, modellando il sistema nervoso con quello che oggi si chiama una rete neurale. Il modello permette di prevedere gli stati futuri del sistema, ma non di recuperare quelli passati, perché l'accesso alla memoria ne modifica il contenuto: cioè, il modello di McCulloch andava in direzione opposta a quella di Freud. Comunque, quando lo incontrai, McCulloch cambiò la mia vita. Ma anche il suo collaboratore Pitts è stato fondamentale per me, perché mi ha insegnato l'importanza dei libri: ne prendeva a palate dalla biblioteca, e li riportava indietro con un carretto solo quando lo minacciavano, ma poi ricominciava».

Quali altri contributi hanno fornito McCulloch e Pitts?

«Il loro lavoro oggi è diventato un classico. Agli inizi i neurofisiologi lo criticarono perché i neuroni artificiali del modello potevano essere sia eccitati che inibiti, mentre quelli naturali del cervello si pensava non venissero mai inibiti. Ma tre anni dopo si scoprirono i neuroni inibitori nel cervello, e si capì che avevano ragione McCulloch e Pitts, e non i neurofisiologi. Tra l'altro, la ragione era semplice: senza inibizione i circuiti neuronali potrebbero calcolare solo funzioni monotone, che salgono sempre o scendono sempre, ma non quelle che a volte salgono e a volte scendono. E questo è appunto il

vantaggio predittivo dei modelli semplici e precisi, a cui ho accennato».

Quale fu il suo passo successivo?

«Nel 1964 feci la mia tesi sulla complessità con Marvin Minsky, uno dei padri dell'Intelligenza Artificiale, che in quel periodo stava scrivendo il suo bel libro sulle *Macchine finite e infinite* e la loro capacità di calcolo. Ricordo ancora che uno degli esercizi di quel libro chiedeva di costruire una macchina in grado di autoriprodursi: la soluzione stava nel costruire una macchina universale capace di costruire macchine arbitrarie attraverso le loro descrizioni, e poi di darle in pasto la propria descrizione».

C'è qualche legame con il suo lavoro attuale sulla coscienza?

«Direi di sì. Ad esempio, mi piacerebbe capire la coscienza in maniera analoga al modo in cui i logici capiscono la calcolabilità: attraverso molte definizioni apparentemente diverse, ciascuna delle quali cattura qualche particolare aspetto della nozione, ma che poi risultano essere tutte equivalenti fra loro. Questo proverebbe, come nel caso della calcolabilità, che ci si è imbattuti in qualcosa di veramente fondamentale».

Non crede che la coscienza lo sia, a priori?

«Non lo so: potrebbe non esserlo, e rivelarsi meno cruciale di quanto ci appaia. Ma quello potrebbe essere un bel modo per dimostrare che è fondamentale».

E quale sarebbe, in questa analogia, l'analogo dell'approccio di Alan Turing alla calcolabilità attraverso i computer?

«Ce ne sono almeno due. Uno è l'approccio di Bernard Baars alla coscienza come un palcoscenico di teatro, sul quale alcuni attori vengono momentaneamente illuminati sulla scena, mentre gli altri attori stanno al buio in platea o dietro le quinte, in attesa del loro

turno. L'altro è l'approccio di Giulio Tononi, che sfrutta invece la metafora della veglia e del sonno. Attualmente queste mi sembrano le uniche teorie neurofisiologiche passibili di una qualche modellizzazione informatica».

Lavora direttamente con questi neurofisiologi?

«No, non li ho mai incontrati.

Anche perché sono interessato a vedere cosa verrebbe fuori dalla mia testa, come se fossi l'unica persona esistente al mondo».

Solo senza altri uomini, o anche senza Dio?

«Fino all'adolescenza mi sono sempre messo in pace con Dio, prima di addormentarmi la sera, ma arrivato all'università ho scoperto che nessuno dei miei

amici credeva. Così ho chiesto a Dio che mi scusasse, se per un mese mi assentavo per cercare di capire come stavano i miei amici senza di lui. E poi, naturalmente, non sono mai tornato indietro. Ma devo stare attento a non farlo anche con gli uomini, perché si può vivere felicemente senza Dio, ma non senza gli altri».

© RIPRODUZIONE RISERVATA

TERZA PAGINA

Vincitore del premio Turing, lo studioso americano indaga sulla natura fisiologica del nostro essere consapevoli e sull'eventualità che possano esserlo anche le macchine

“
La mia disciplina ha il grande vantaggio di fornire definizioni precise, verificabili o refutabili, ma sempre in maniera sperimentale
”



Scienziato

Manuel Blum è nato a Caracas nel 1938 e successivamente è diventato cittadino americano. È un informatico e nel

1995 ha ricevuto il premio intitolato ad Alan Turing per il suo contributo nel campo della teoria della complessità computazionale e della crittografia

