

UOMINI E MACCHINE

Un libro mette a nudo, su basi scientifiche, i limiti dell'Intelligenza artificiale

Computer, cosa sai fare?

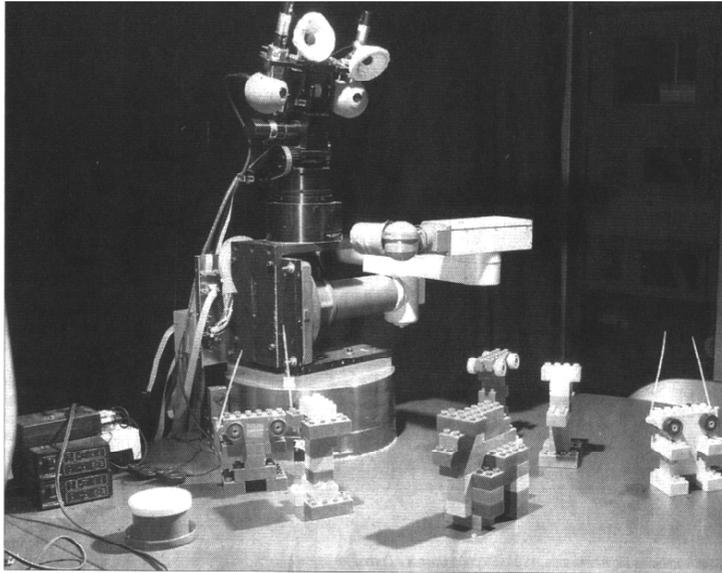
DI DIEGO MARCONI

Nel 1972, il filosofo californiano Hubert Dreyfus pubblicò un libro intitolato *Quel che i computer non sanno fare* (trad. it. Armando 1989). Dreyfus sosteneva che l'Intelligenza Artificiale — che in quegli anni coglieva successi significativi, e sembrava destinata a grandi e rapidi progressi — era la realizzazione pratica delle idee filosofiche del razionalismo europeo, ne condivideva l'immagine della mente umana e sarebbe fallita per le stesse ragioni per cui era filosoficamente fallito il razionalismo: e cioè perché la mente, con buona pace di Cartesio, Hobbes e Leibniz, non è essenzialmente elaborazione di rappresentazioni secondo regole. Il libro ebbe molta fortuna, forse perché leniva una duplice ansia: da un lato, se Dreyfus aveva ragione una mente artificiale era impossibile e dunque non rischiamo di essere "soppiantati dalle macchine"; dall'altro, la dimensione spirituale dell'umanità era messa in salvo dal rischio della riduzione a «nient'altro che programmi di computer». Nella comunità filosofica, peraltro, le tesi del libro erano e restarono controverse. Pochi pensarono che Dreyfus avesse dimostrato una volta per tutte che i computer non saranno mai capaci di riconoscere un volto o di mettere ordine in una stanza.

Il libro di David Harel, appena tradotto ottimamente da Luigi Civalleri per Einaudi, porta come sottotitolo, nell'edizione originale, «Quel che i computer davvero non sanno fare». I limiti del computer di cui si tratta nel libro, infatti, non sono argomentati filosoficamente come nel libro di Dreyfus, ma risultano da dimostrazioni matematiche; e perciò siamo davvero sicuri — se lo siamo di qualcosa — che quei limiti non possono e non potranno mai essere superati. Tuttavia, gli umanisti nemici delle macchine non hanno motivo di cantare vittoria, perché questi problemi, che il computer non sa risolvere,

L'impossibilità di risolvere certi problemi è dimostrata matematicamente. E riguarda anche gli esseri umani

Babybot, il robot umanoide realizzato presso il LIRA-Lab dell'Università di Genova con l'obiettivo di studiare i meccanismi dell'apprendimento e dello sviluppo dell'intelligenza nell'uomo. Babybot impara a «conoscere il suo corpo» e a controllarne il movimento sulla base delle informazioni sensoriali provenienti dagli occhi, dalle orecchie, dall'organo dell'equilibrio (vestibolare) e dai sensori racchiusi nei muscoli (motori) e nelle articolazioni (giunti e snodi). L'apprendimento avviene misurando e correggendo gradualmente gli errori di esecuzione



non sono nemmeno alla portata della mente umana. Quel che i computer davvero non sanno fare, non lo sappiamo fare neanche noi.

Semplificando un po', si tratta qui di due tipi di limiti. Ci sono anzitutto problemi *indecidibili*, per cui cioè non esiste e non può esistere un procedimento (algoritmo) di risoluzione. Per esempio, non è possibile scrivere un programma di computer che stabilisca se un qualsiasi altro programma, dato un certo *input*, fornisce un risultato per quell'*input* e si ferma, o invece continua a girare indefinitamente. Questo problema, detto "della fermata", è indecidibile. E ce ne sono tanti altri. Molti di essi sono equivalenti, nel senso che se uno di essi fosse decidibile — se ci fosse un oracolo che lo decide — allora potremmo usare quello stesso oracolo per decidere anche gli altri. Ma altri non sono così. Prendete, ad esempio, il sogno di ogni program-

matore: un algoritmo che controlla i programmi, nel senso che, dato un programma, risponde «Sì» se il programma termina e fornisce la soluzione corretta per tutti gli *input* ammissibili, e «No» se per qualche *input* o non termina, o fornisce una soluzione scorretta. Il problema del controllo dei programmi è indecidibile; non solo: resterebbe indecidibile anche se disponessimo di un oracolo per il problema della fermata o per uno degli altri a esso problemi equivalenti. Ci sono quindi vari gradi di indecidibilità: quello più alto è rappresentato dai problemi — detti *fortemente indecidibili* — che non sarebbero risolti nemmeno disponendo di una gerarchia infinita di oracoli.

L'altro tipo di limite è qualitativamente diverso, ma in pratica non meno importante. Ci sono problemi per cui esistono algoritmi di risoluzione — si tratta cioè di problemi decidibili — ma questi algo-

ritmi sono troppo *lenti*: al crescere della lunghezza dell'*input*, il tempo della loro esecuzione cresce al di là del tollerabile. Per esempio, per un *input* di lunghezza 100 forniscono sì la risposta, ma ci mettono un numero di secoli di 185 cifre (si ricordi che dal *Big Bang* a oggi è trascorso un numero di secoli di 9 cifre). Si dirà: ma i computer non diventano sempre più veloci? Certo; ma questo vuol solo dire che, se i computer diventassero un milione di volte più veloci di quello che sono oggi, un algoritmo come quello di cui stiamo parlando darebbe la risposta dopo un numero di secoli di 178 cifre, anziché 185. Esiguo progresso.

I problemi che sono decidibili soltanto da algoritmi di questo genere (detti *esponenziali*, perché se N è la lunghezza dell'*input* il tempo di esecuzione è m^N) si chiamano "intrattabili". Non si deve pensare che si tratti di sofisticate que-

sti, di reti di comunicazione eccetera. Anche di questo problema non sappiamo se sia trattabile: anche qui, gli algoritmi di risoluzione noti sono lentissimi.

La questione dell'esistenza di algoritmi che risolvano questo tipo di problemi in un tempo ragionevole (detto tempo *polinomiale*, perché dipende da N^n , dove N è la lunghezza dell'*input*) è uno dei grandi problemi aperti della matematica attuale. Sappiamo che, se esistono algoritmi "ragionevoli" per uno di questi problemi, allora esistono per tutti: sorprendentemente, questi problemi sono tutti equivalenti. Ma non sappiamo se questi algoritmi esistano (anche se la maggior parte dei matematici pensa di no). Nella pratica, dobbiamo allora ricorrere ad algoritmi approssimanti, che garantiscono di fornire la soluzione corretta con una probabilità molto elevata, ma sempre minore di 1. Questi algoritmi — davvero geniali, e presentati da Harel in modo affascinante — esistono per alcuni problemi ma non per altri: fino a poco tempo fa si pensava che dovessero esistere per tutti, ma, a quanto pare, non è così.

Questo libro, scritto da un *computer scientist* di primo piano che è anche un grande divulgatore, ha il raro merito di mettere alla portata di chiunque sappia leggere l'italiano problemi scientifici che si collocano alla frontiera della ricerca. Ma ha anche altri due pregi. In primo luogo, è un'ottima introduzione ad alcuni concetti centrali dell'informatica: chiunque voglia sapere che cosa fa un computer (elettronica a parte) dovrebbe leggerlo. In secondo luogo, propone contenuti scientifici in modo accattivante e a tratti anche divertente, ma senza enfasi e senza presentare quei particolari contenuti come la chiave della comprensione dell'intero universo; in questo genere di letteratura, non capita tutti i giorni.

David Harel, «Computer a responsabilità limitata. Dove le macchine non riescono ad arrivare», traduzione di Luigi Civalleri, Einaudi, Torino 2002, pagg. 196, € 13,00.