

## L'INTELLIGENZA E L'ARTIFICIALE

Nel 1774, pochi anni dopo la clamorosa pubblicazione de *L'homme-machine* di La Metrie, un giovane di nome Pierre Jacquet-Droz presentò a La Chaux-de-Fonds tre splendidi automi in figura umana capaci rispettivamente di disegnare, scrivere, ed eseguire su un piccolo clavicembalo interi brani musicali respirando regolarmente e seguendo lo spartito con gli occhi. Possiamo immaginare la reazione del pubblico. L'antico sogno umano di essere non solo *creature* ma anche *creatori* sembrava avverarsi. E ciò che più colpiva era la natura delle prestazioni di questi “androidi”, in grado non solo di imitare comportamenti *fisici* tipici dell'uomo, ma addirittura di simulare attività *intellettuali* quali, appunto, disegnare, scrivere o suonare.

Naturalmente non si può dire che le creature di Jacquet-Droz fossero davvero dotate di “intelletto”. Capolavori di tecnologia orologiaia, automatismi di squisita fattura, queste macchine eseguivano solo dei programmi meccanici prestabiliti. Non avevano capacità proprie. Erano, in sostanza, dei grossi (ancorché complicatissimi) *carillon*.

Semmai si sarebbe potuto dire diversamente qualche decennio più tardi, quando si diffuse la notizia di un automa capace persino di giocare a scacchi. A differenza dei tre androidi, il “Turco” creato da Wolfgang von Kempelen (lo stesso cui si devono le celebri fontane idrauliche di Schönbrunn) agiva dinamicamente, decidendo di volta in volta le mosse da effettuare a seconda delle mosse dell'avversario. E lo faceva così bene che, pare, riuscì addirittura a battere Napoleone.

Peccato che anche in questo caso l'apparenza superasse la realtà. A quanto si dice, la meravigliosa creatura di von Kempelen era ... guidata da un operatore (ben nascosto agli astanti), e il destino volle che un incendio la distruggesse nel bel mezzo di una sfida.

Ma che dire del “Motore Analitico” di Charles Babbage? Sfruttando l’idea del telaio di Jacquard, questa macchina doveva eseguire calcoli matematici sotto il controllo di una scheda perforata. E l’idea parve talmente promettente che la figlia di Lord Byron, Ada Augusta Lovelace, giunse a scrivere nel suo diario che una tale macchina avrebbe potuto elaborare non solo numeri, ma qualunque informazione fosse riducibile a un insieme di simboli – anticipando così l’idea del computer digitale.

Ironia della sorte, il Motore Analitico non venne mai costruito. E il nome di Babbage fece eco nelle cronache mondane del tempo più per la pickwickiana battaglia che egli condusse contro l’inquinamento acustico dei suonatori di organetto che per le sue intuizioni scientifiche.

**I**n questo scenario di prodigi apparenti e progetti incompiuti, è solo in tempi recenti che il vecchio sogno di creare macchine in qualche modo dotate di “intelletto” ha acquisito un significato più preciso.

Nel 1950 il logico inglese Alan Turing pubblicò un articolo in cui poneva il problema in termini essenzialmente operazionali, suggerendo di analizzarlo mediante una sorta di *gioco d’imitazione*. Le regole: un esaminatore pone delle domande a un uomo e a una macchina nascosti in una stanza. Se dalle risposte egli riesce a distinguere l’uomo dalla macchina, allora vince l’uomo. Se invece non ci riesce, allora la macchina vince e si merita l’attributo di “intelligente”.

Evidentemente, né gli androidi di Jacquet-Droz né i rozzi automi dell’800 avrebbero potuto competere. Tuttavia lo stato dell’arte degli anni cinquanta era profondamente diverso. Non solo perché l’avvento dei computer digitali aveva in qualche modo dato corpo (hardware) alle idee di Babbage e alle profezie di Lady Lovelace. Ma anche perché nel frattempo si erano registrati importanti progressi teorici. La nozione intuitiva di computabilità effettiva aveva ceduto il posto alla nozione rigorosa di computabilità *ricorsiva*. E soprattutto, si era dimostrato che ogni funzione ricorsivamente computabile può essere com-

putata in un tempo finito da una macchina guidata da regole ricorsivamente applicabili. Orbene, i computer digitali *sono* macchine di questo tipo: quindi, ammesso che la funzione che produce connessioni di input-output intelligenti in una persona sia effettiva, un comune computer potrebbe in linea di principio computarla, superando il test proposto da Turing. L'unico problema è identificare tale funzione e stilare il programma che la realizzi. Un problema tutt'altro che banale, beninteso, ma sufficientemente definito sul piano teorico: la soluzione è una questione puramente operativa.

**È** dunque possibile realizzare macchine “intelligenti”? Sicuramente l'articolo di Turing segna la nascita dell'Intelligenza Artificiale (IA) come disciplina *scientifica*, ben diversa dai goffi tentativi che l'avevano preceduta. Ma è davvero possibile riuscire nell'intento implementando programmi in un computer digitale?

Forse l'obiezione più radicale è quella formulata in tempi recenti dal filosofo John Searle. Immaginiamo – dice Searle – che un italiano che non capisce una parola di cinese sia chiuso in una stanza in cui vi sono dei cestelli colmi di ideogrammi cinesi. Supponiamo anche che nella stanza vi sia un librone scritto in italiano che spiega come manipolare e trascrivere questi simboli con altri simboli. Infine, immaginiamo che da una fessura entrino di tanto in tanto dei foglietti con degli ideogrammi: il nostro signore li prende, consulta il librone, traffica un po' con i cestelli, trasforma i simboli e ricopia il risultato su un altro foglietto che poi spedisce all'esterno.

La metafora è chiara: il tizio nella stanza è un computer, il librone è un programma, e i cestelli colmi di ideogrammi forniscono la base dati. Ebbene – incalza Searle – se all'esterno della stanza c'è un tizio cinese che trova ragionevoli le “risposte” che provengono dall'interno, egli potrà anche pensare di interloquire con un cinese: in altre parole, il computer supererà il test di Turing per la conoscenza del cinese. Ma evidentemente la verità è un'altra: il tizio all'interno (i.e. il computer) non fa che manipolare scarabocchi per lui incomprensibili. Una

operazione puramente *sintattica*, mentre l'intelligenza ha contenuto *semantico*.

È possibile replicare? Qualcuno ha osservato come l'argomento di Searle sia reminiscente di un pattern poco edificante nella storia del pensiero scientifico-filosofico. Berkeley trovava inaccettabile che il suono fosse (riducibile a) puri e semplici fenomeni ondulatori, ma la ricerca dimostrò esattamente il contrario. Oppure pensiamo al coro di critiche che investì l'ipotesi di Maxwell sulla natura della luce. «La luce non può essere un fenomeno elettromagnetico, altrimenti basterebbe agitare un magnete in una stanza buia per illuminarla!». Non è questo, in fondo, il genere di argomento esemplificato della “stanza cinese”? Il tizio italiano non comprende il significato delle proprie manipolazioni simboliche. Ma il manuale contiene per ipotesi tutte le regole necessarie perché tali manipolazioni risultino ragionevoli, sensate, dotate di significato. La semantica è lì. Perché allora negare che il manuale *più* il tizio che lo usa – cioè il programma *più* il computer – conoscono il cinese?

Qualcuno ha anche obiettato che l'argomento di Searle è in realtà una critica al test di Turing, piuttosto che all'IA. Ben venga: dopo tutto il test è stato il punto di partenza ma non per questo deve costituire un rigido criterio di riferimento. Il fatto è che l'argomento di Searle (a differenza del test di Turing) fa appello a principi non comportamentali: se il *Gedankenexperiment* della stanza cinese suggerisce le conclusioni volute, è perché noi *sappiamo* che il tizio all'interno della stanza non conosce il cinese. Ma questa è un'arma a doppio taglio. Scetticismo per scetticismo, cosa esclude che un osservatore neutrale – un dio epistemologico, per così dire – possa affermare lo stesso del tizio all'*esterno*?

E poi il test di Turing è un test *globale*, mentre l'argomento di Searle presuppone che il computer superi il test *relativamente* alla conoscenza della lingua cinese. Qual è la linea di demarcazione fra questa e le altre capacità mentali? Che ragioni vi sono per supporre che un sistema, artificiale o non, possa conoscere il cinese *senza sapere nulla di tutto il resto*?

**D**omande difficili, alle quali forse non è possibile dare una risposta senza in qualche modo riproporre il quesito di fondo. E forse non è senza una certa presunzione che si può pensare di risolverlo semplicemente esaminando posizioni filosofiche contrastanti.

In effetti oggi l'IA si configura sempre più come una scienza sperimentale, pragmatica prima ancora che teorica. È l'obiettivo pratico di realizzare sistemi dotati di capacità comportamentali sempre più elevate che sottende i principali centri di ricerca in questo settore, primo fra tutti l'IRST di Trento. Per Luigi Stringa, ad esempio, l'aspetto filosofico potrebbe anche non essere determinante: con cinque miliardi di esseri umani più o meno pensanti in circolazione, poco importa che anche le macchine *pensino* davvero, o *siano* davvero intelligenti. Quel che conta è sviluppare sistemi che si *comportino* come tali, così da poterci dare una mano quando serve. E da questo punto di vista lo storico “computo, ergo sum” da cui è partita l'IA è solo una delle possibili strategie di ricerca. Anzi, forse la computazione non è poi così importante, e Stringa non esclude che in molti casi la giusta via al comportamento intelligente risieda nella memoria piuttosto che nella capacità di elaborazione.

Un discorso lungo, che non può finire qui. In fondo l'Intelligenza Artificiale è solo agli inizi mentre quella naturale può già vantare una certa età. Diciamo che la ricerca ci darà macchine sempre più evolute. E prima di dare un voto alla loro intelligenza, non sarebbe una brutta idea scambiare due chiacchiere con loro e ascoltare con attenzione ciò che esse avranno da dirci.