Kevin Warwick

INTELLIGENZA ARTIFICIALE

LE BASI

Prefazione di *Carlo Alberto Avizzano*

WEBBOOK

#saggitudine

SCOPRI DA DOVE VIENE L'A.I. E IN CHE DIREZIONE STA ANDANDO

CAMBIARE CAPPELLO SIGNIFICA CAMBIARE IDEE,
AVERE UN'ALTRA VISIONE DEL MONDO.

C.G. Jung

Dario Flaccovio Editore

Kevin Warwick

INTELLIGENZA ARTIFICIALE
LE BASI

Traduzione di
Chiara Barattieri di San Pietro
Giuseppe Maugeri

Prefazione di
Carlo Alberto Avizzano

WEBBOOK
#saggitudine

Titolo originale: *Artificial Intelligence - The Basics*

Per l'edizione italiana:
KEVIN WARWICK
INTELLIGENZA ARTIFICIALE - LE BASI
ISBN 978-88-579-0432-0

Traduzione: **Chiara Barattieri di San Pietro** e **Giuseppe Maugeri**

Warwick, Kevin <1954->

Intelligenza artificiale : le basi / Kevin Warwick ; traduzione di Chiara Barattieri di San Pietro, Giuseppe Maugeri ; prefazione di Carlo Alberto Avizzano. – Palermo : D. Flaccovio, 2015.
(#saggitudine)
ISBN 978-88-579-0432-0
1. Intelligenza artificiale. I. Barattieri di San Pietro, Chiara.
II. Maugeri, Giuseppe.
006.3 CDD-22 SBN PAL02880758

CIP - Biblioteca centrale della Regione siciliana "Alberto Bombace"

Prima edizione: giugno 2015
Stampa: Tipografia Priulla, Palermo, giugno 2015

webintesta.it Copertina: Illustrazione realizzata da Goran Factory

La traduzione dell'opera è stata realizzata grazie al contributo del SEPS
SEGRETARIATO EUROPEO PER LE PUBBLICAZIONI SCIENTIFICHE

Via Val d'Aposa 7 - 40123 Bologna - seps@seps.it - www.seps.it

Indice

Presentazione

Generare macchine simili all'uomo è sempre stata una delle principali ambizioni della ricerca e dello sviluppo sin dalle origini della rivoluzione industriale. I primi sistemi che emulavano il comportamento umano, durante l'esecuzione di particolari attività, furono soprannominati "automi", ed erano, in sostanza, giocattoli meccanici di precisione governati da meccanismi a orologeria in grado di generare complesse sequenze di movimento. All'inizio del 1700 centinaia di questi sistemi cominciarono a essere prodotti per essere esibiti presso le corti reali o a pagamento in esposizioni dedicate. Primi tra questi il pianista, il disegnatore e lo scrivano di Jacquet-Droz e fratelli, prodotti per la corte di Luigi XVI. L'ultimo, un automa in grado di scrivere frasi arbitrarie lunghe fino a 40 lettere.

Macchine come il flautista di Jaques Vaucanson prodotto nel 1738, pur essendo degli artefatti di nessuna utilità pratica, manifestavano il desiderio di alcuni uomini e della società in generale di veder nascere macchine non solo dall'aspetto umano ma, soprattutto, capaci di imitare l'uomo nelle sue tipiche abilità gestuali e intellettuali.
Già nel 1770 il barone von Kempelen, stimolato a produrre dalla corte austriaca una macchina che potesse impressionare l'imperatrice Maria Teresa, si cimentò nella realizzazione di quello che è comunemente noto come il giocatore di scacchi. Un manichino seduto di fronte a una scacchiera in grado di giocare autonomamente al più nobile dei giochi e battere in questo i migliori giocatori di corte.
Una vera macchina intelligente, frutto di una combinazione di meccanismi impressionanti, e certamente dell'inganno, visto che la vera intelligenza del dispositivo era delegata ad un abile scacchista umano in grado di nascondersi nel cuore della macchina e di pilotare gli organi di governo dell'automa per gareggiare con l'ignaro sfidante.
Oggi, a oltre tre secoli di distanza, molte cose sono cambiate, ma non la voglia originaria di generare macchine sempre più potenti, capaci di rassomigliarci e/o sostituirci nei compiti meno gratificanti. Di certo, le opere di inizio secolo hanno avuto un ruolo fondamentale per la rinascita e il rinvigorimento di interesse per questo tipo di sistemi. Ricordiamo tra queste R.U.R. (Rossum's Universal Robots) di Karel Čapek, in grado di coniare e dare vita al concetto di Robotica, e la produzione letteraria di Isaac Asimov che nella sua opera *Runaround* fu in grado di unire alla propria fantasia un'etica comportamentale per questi sistemi, costituita da un insieme di regole rigide e inviolabili, a oggi ancora riconosciute come le leggi fondamentali della Robotica.

Di lì a poco, anche grazie alla nascita dell'informatica e del calcolo computazionale, le macchine pensanti sono divenute una realtà. Sicuramente una data importante è il 1956, quando Allen Newell e Herb Simon, al ritorno dalla pausa natalizia, annunciarono ai propri studenti di aver costruito una prima macchina pensante, in grado cioè di utilizzare un insieme di regole formali per risolvere autonomamente problemi descritti in forma matematica. Fu nello stesso anno che John McMarty coniò in un seminario interdisciplinare il termine di "Artificial Intelligence", ripreso in seguito per indicare una nuova disciplina di ricerca volta a far compiere alle macchine ragionamenti e operazioni che naturalmente richiederebbero un'intelligenza umana.

Kevin Warwick è professore di cibernetica presso l'università di Reading; il suo libro "Artificial Intelligence – Le basi" ripercorre un avvincente e interessante viaggio tra le tecniche per la realizzazione di macchine intelligenti, sintetizzando i risultati raggiunti partendo dalla base delle fondamenta gettate da matematici, artisti e filosofi a inizio secolo, e il meccanismo di funzionamento degli algoritmi sviluppati da molteplici ricercatori a partire dagli anni sessanta.

Il libro di Warwick fornisce al lettore una panoramica molto vasta sulle tecniche oggetto della ricerca, sul significato dei risultati raggiunti e sulle implicazioni all'interno della nostra società. Tra i molteplici libri disponibili sul tema, il lavoro di Warwick offre sicuramente una prospettiva originale e interessante che merita di essere letta e ragionata.

Rispetto a diversi classici accademici, l'autore predilige un approccio prevalentemente divulgativo: linee di ricerca, anche molto complesse, vengono rese accessibili al grande pubblico in modo chiaro, sintetico e facilmente comprensibile; la complessità e il rigore degli strumenti matematici utilizzati vengono semplificati a favore della chiarezza

espositiva, senza tuttavia per questo alterare nel significato il cuore computazionale di ciascun algoritmo.
L'autore tuttavia non si limita a raccontare algoritmi e tecniche dalla prospettiva di un osservatore esterno, ma per ciascuno di essi offre una propria interpretazione da esperto navigato, che superando nel merito le prestazioni degli algoritmi affronta più dettagliatamente gli aspetti etici e filosofici, e il graduale avvicinamento di tali risultati alle prestazioni umane.
Si capisce pertanto come l'autore non intenda esimersi dal confrontare le prospettive aperte da questi algoritmi nascondendosi dietro gli aspetti meramente numerici dei risultati ottenibili tramite applicazioni sperimentali.
Al contrario, proprio quando le aspettative promesse dalle tecniche di intelligenza artificiale si fanno più ambiziose e con esse lo scetticismo scientifico, l'autore approfondisce maggiormente l'interpretazione dei risultati raggiunti.
Warwick affronta per ciascuno di essi il tema della "manifesta superiorità umana" perorato da diverse parti e teso a meccanicizzare il ruolo della "intelligenza artificiale" sminuendola in posizioni subalterne alla "intelligenza umana", quasi quest'ultima dovesse sempre rimanere dotata di caratteristiche uniche, mai replicabili in formule e algoritmi.
In questo contesto, Warwick non si limita a presentare le tecniche attualmente esistenti e i risultati raggiungibili ma per ciascuna di esse desidera argomentare se la superiorità dell'intelligenza umana sia reale o meno.
Per comprendere appieno il contesto in cui nasce il desiderio dell'autore di rimettere in discussione un confronto oggettivo tra quanto ottenibile oggi con l'intelligenza artificiale e le facoltà espresse dall'intelligenza umana, è probabilmente necessario interrogarsi sulle motivazioni che portano alla presenza diffusa di uno scetticismo nei confronti

dell'intelligenza artificiale. Un fenomeno simile è noto ed è stato documentato come "uncanny valley".
La "uncanny valley" è un fenomeno psicologico, presentato da Masahiro Mori nel 1970, e relativo all'accettabilità estetica di figure e comportamenti. L'ipotesi di Mori è che il senso di familiarità ed accettazione (empatia) di alcune entità che emulano la figura umana cresce quanto più aumenta la rassomiglianza tra le entità e le caratteristiche umane stesse. Esiste però un intervallo di rassomiglianza negativo tale per cui, se la verosimiglianza supera una data soglia, l'individuo implicitamente rifiuta ed interpreta come distorsione tutto ciò che gli viene presentato.
Tali sensazioni negative di repulsione e turbamento persistono fino al superamento di una seconda soglia di verosimiglianza oltre la quale l'empatia cresce molto rapidamente. In estetica l'uncanny valley viene ad esempio associata alla figura dei mostri e degli zombie, che pure ricordando l'estetica umana, producono certamente più repulsione di cartoni animati o altre stilizzazioni del comportamento dell'individuo che siamo abituati ad accettare.
Warwick affronta pertanto le posizioni di scetticismo cercando dei confronti oggettivi con i risultati raggiunti dalla ricerca negli ultimi trent'anni. I capitoli centrali del libro prendono spunto dalla revisione di alcune delle competizioni esistenti tra uomo e robot, mirate ad appurare le capacità di algoritmi di intelligenza artificiale di replicare caratteristiche tipicamente umane, quali: le capacità dialettiche (originariamente proposte nel test di Turing); l'abilità nel gioco e nella strategia; le capacità di evolversi autonomamente in base all'esperienza e all'ambiente. Le conclusioni dell'autore sono tipicamente in controtendenza rispetto alle posizioni scettiche. Secondo Warwick, nonostante in alcuni confronti i risultati rimangano ancora a favore dell'uomo, in

molte delle competizioni attuali le macchine hanno ampiamente superato le potenzialità umane (ad esempio gli scacchi), oppure sono vicine a un confronto quantomeno equalitario.
Le posizioni dell'autore sono spesso molto forti, e ne emerge un libro decisamente intrigante ed interessante, in grado di trasformare quello che altrimenti sarebbe un glossario di tecniche in uno scontro di altissimo livello tra aspettative e risultati.
Il lavoro di Warwick offre sotto questa prospettiva un'interpretazione nuova dei progressi recenti percorsi dall'intelligenza artificiale, lontana da eventuali condizionamenti psicologici che potrebbero essere causati dalla presenza di una uncanny valley. Non solo, dopo aver rappresentato le principali tecniche di analisi e progettazione allo stato attuale dell'intelligenza artificiale, l'autore rielabora le potenzialità di questi algoritmi sulla base dei nuovi progressi tecnologici, che non si limitano ad offrire un cervello ma, grazie ad aumentate capacità di sensing, provvedono alle macchine un corpo in grado di popolare ed orientare, con una grossa quantità di informazioni ambientali, i ragionamenti argomentativi dell'intelligenza artificiale.
Con il suo libro Warwick intende proiettarci nel futuro – ben conscio che buona parte del cervello umano è dedicata all'interpretazione dei segnali provenienti dal corpo piuttosto che al puro ragionamento – reinterpretando in chiave digitale la linea di pensiero filosofica nota come "embodied cognition" e proposta in informatica da ricercatori come Mark Greenlee (Neuroscienze), Rodney Brooks (Robotica), o Rolf Pfeifer (Intelligenza Artificiale).
Rispetto alle numerose pubblicazioni esistenti sull'argomento, il lavoro di Warwick si distingue per i suoi temi caratteristici di interpretazione e prospezione, e la sua let-

tura risulta gradevole sia a chi, totalmente avulso dai temi di ricerca, intende avere un'introduzione gentile e panoramica alla ricerca del settore, sia a chi, già indottrinato degli algoritmi, intenda prendersi un momento di riflessione e di interpretazione su quali siano le tesi e le antitesi poste a supporto e a confutazione dei risultati finora raggiunti. Warwick ci offre la sua posizione, a favore di una difesa quasi paterna del valore di tali risultati. Alcune sue affermazioni (come il tema degli alieni) potranno sembrarvi quasi provocatorie, ma proprio per questo risulteranno fonte di ulteriori riflessioni. Può il figlio superare il padre? Può l'allievo superare il maestro? Potranno questi algoritmi superare la nostra capacità di ragionare in un futuro non molto lontano? Lo hanno già fatto e non ce ne siamo accorti? Pur essendo solo un libro di base, il saggio di Warwick è illuminante per esplorare le implicazioni di queste ipotesi sotto molteplici punti di vista.

Carlo Alberto Avizzano
Pisa, 11 gennaio 2015

Prefazione

Il campo dell'Intelligenza Artificiale (AI, Artificial Intelligence) vide concretamente la luce tra gli anni '40 e '50 del secolo scorso, in concomitanza con la nascita dei computer. Le fasi iniziali del suo sviluppo erano decisamente focalizzate su come far sì che i computer svolgessero operazioni che, eseguite da un essere umano, sarebbero state considerate intelligenti. Ciò si traduceva essenzialmente nel tentativo di fare in modo che i computer imitassero gli esseri umani in alcuni o in tutti gli aspetti del loro comportamento. Negli anni '60 e '70 questo aprì un dibattito filosofico su quanto i computer potessero avvicinarsi al cervello umano, e sulla reale rilevanza delle eventuali differenze. Tale periodo – che in questo libro chiameremo dell'"AI classica" – espresse tuttavia un potenziale piuttosto limitato.

Gli anni ’80 e ’90 videro l’affermazione di un metodo completamente nuovo con cui affrontare il problema, una sorta di approccio bottom-up che mirava al conseguimento dell’AI tramite la concreta realizzazione di cervelli artificiali. Questo aprì possibilità radicalmente nuove, sollevando al tempo stesso una nuova serie di questioni. L’AI non era più ristretta alla pura e semplice imitazione dell’intelligenza umana: adesso poteva essere intelligente in sé e per sé.

Nonostante in certi casi potesse essere ancora ottenuta replicando le procedure attraverso cui opera un cervello umano, aveva ormai il potenziale per essere più estesa, più veloce e migliore. Dal punto di vista filosofico ne conseguiva che finalmente un cervello artificiale poteva virtualmente raggiungere prestazioni superiori a quelle di un cervello umano.

In anni più recenti il settore ha registrato un notevole incremento. Applicazioni AI nel mondo reale, in particolare in ambito finanziario, produttivo e militare, operano in modalità con le quali, semplicemente, il cervello umano non può competere. I cervelli artificiali sono sempre più spesso dotati di un corpo attraverso cui percepire il mondo e muoversi nello spazio per modificarlo. Sono dotati di capacità di apprendimento e adattamento, e sono in grado di portare a termine quanto si prefiggono nei confronti degli esseri umani. Cosa, questa, che solleva ogni genere di interrogativo in proiezione futura.

Questo libro si propone l’obiettivo di offrire una panoramica davvero moderna e attuale sul campo dell’AI nella sua interezza. Non si tralascerà di esaminare l’AI classica, ovviamente, ma solo in quanto parte di un ambito più complesso. Ed è con il medesimo criterio che si guarderà all’AI moderna. In particolare, saranno discusse alcune tra le più

recenti ricerche sull'AI *embodied*[1] e sull'AI biologica in vitro. L'intento è quello di fornire una guida essenziale e comprensibile sul campo dell'AI oggi, che consenta di capire da dove viene e quale direzione potrebbe prendere. Sebbene il suo principale obiettivo rimanga quello di introdurre all'argomento quanti vi si accostano per la prima volta, l'opera potrebbe rivestire anche un certo interesse per coloro che, pur essendo già impegnati nel campo della scienza, della tecnologia e perfino dell'informatica, hanno magari bisogno di rimettersi al passo con gli sviluppi più recenti.

Vorrei ringraziare molte persone per avermi aiutato a mettere insieme questo libro. In particolare, i miei colleghi e i ricercatori dell'Università di Reading, soprattutto Mark Gasson, Ben Hutt, Iain Goodhew, Jim Wyatt, Huma Shah e Carole Leppard, per il loro contributo significativo. Desidero anche estendere la mia gratitudine a Andy Humphries della Taylor & Francis, che mi ha spinto a ultimare il lavoro, nonostante i miei numerosi impegni e il poco tempo a disposizione. Infine, vorrei ringraziare mia moglie Irena per la sua pazienza e i miei ragazzi, Maddi e James, per le loro critiche.

Kevin Warwick
Reading, gennaio 2011

[1] Vedi nota 4

Introduzione

SINOSSI

Nel capitolo iniziale sarà data una breve visione d'insieme sull'argomento del libro, sui suoi obiettivi e sui potenziali lettori a cui è rivolto. Si ripercorreranno altresì sinteticamente gli sviluppi che quest'area disciplinare ha attraversato nel corso degli anni, citando le figure chiave che ne hanno costellato il percorso, nonché le questioni e le svolte importanti. In sostanza, il capitolo si propone di introdurre in maniera graduale i lettori alla materia. Non è dunque indispensabile per chi abbia già familiarità con l'argomento; tuttavia, potrebbe stimolare qualche riflessione o fornire utili spunti informativi.

PREMESSA

Concepito come un manuale introduttivo al campo dell'In-

telligenza Artificiale (AI), il libro si propone di fornire materiale a un primo corso per studenti che intendano specializzarsi in aree quali informatica, ingegneria e cibernetica. Tuttavia, può essere utile anche come testo generico o di riferimento per tutti gli studenti interessati, in particolare per quelli di altre branche della scienza e della tecnologia.
E, ancora, come libro introduttivo per le scuole secondarie e per i lettori in genere che desiderino farsi un quadro generale del settore in una forma facilmente comprensibile.
Dal momento che negli ultimi anni la materia è andata incontro a cambiamenti spettacolari, questo manuale intende proporre una visione moderna dell'argomento.
Naturalmente, sarà data importanza alle tecniche di AI classica, ma entro certi limiti: l'obiettivo è infatti un testo che sia attuale e onnicomprensivo.
Il libro affronta aspetti dell'AI che chiamano in causa filosofia, tecnologia e metodi di base. Sebbene fornisca indicatori di programmazione AI nelle sue linee fondamentali, non si occupa nel dettaglio della scrittura di programmi reali e non si lascia invischiare nelle complicazioni relative alle differenze tra i linguaggi di programmazione. Il suo scopo primario è quello di delineare una panoramica sull'AI e di costituire una guida essenziale che non vada troppo in profondità su uno specifico argomento. Non mancano i rimandi a ulteriori testi tramite cui il lettore potrà approfondire una particolare area di interesse.
Per quanto la panoramica offerta sia, appunto, d'insieme e potenzialmente accessibile a un pubblico ampio, l'opera è stata scritta con rigore accademico. Se alcuni manuali pubblicati in precedenza strizzavano l'occhio ai libri per i bambini, questo di certo non lo fa.

I primordi dell'AI

Esistono forti legami tra lo sviluppo dei computer e l'affermazione dell'AI. Tuttavia, i semi della disciplina erano già stati piantati ben prima che i moderni computer prendessero forma. Filosofi come Cartesio consideravano gli animali in termini di prestazioni meccaniche, e i cosiddetti automi non erano altro che i precursori degli odierni robot umanoidi. A ogni modo, è possibile retrodatare ulteriormente il concetto di essere artificiale, magari a storie come quella del Golem di Praga o, ancora prima, a miti greci come quello che racconta di Pigmalione e Galatea.

Le radici dirette più salde risalgono probabilmente al lavoro di McCulloch e Pitts che, nel 1943, descrissero modelli matematici (chiamati percettroni) di neuroni (ovvero di cellule cerebrali) ricavati sulla base di un'analisi dettagliata degli originali biologici. I due non solo spiegarono il modo in cui i neuroni si accendono o meno (sono cioè in posizione "on" o "off"), così da operare secondo una commutazione binaria, ma dimostrarono anche come tali neuroni possano apprendere e quindi modificare la loro azione nel tempo.

Uno dei più grandi pionieri in questo campo fu forse lo scienziato britannico Alan Turing. Negli anni '50 (molto prima dell'avvento dei computer come li conosciamo oggi), Turing scrisse un articolo fondamentale in cui tentava di rispondere alla domanda: "Una macchina può pensare?". Se il solo fatto di porsi un simile quesito era, all'epoca, rivoluzionario, inventarsi una prova applicabile (comunemente nota come test di Turing) con cui rispondere alla domanda fu estremamente audace. Il test è analizzato in dettaglio nel Capitolo 3.

Poco tempo dopo, Marvin Minsky e Dean Edmonds realizzarono quello che potrebbe essere descritto come il primo

computer AI, basato su una rete di modelli neurali descritti da McCulloch e Pitts. Negli stessi anni, Claude Shannon considerò la possibilità di un computer in grado di giocare a scacchi e il tipo di strategie necessarie a decidere di volta in volta la mossa successiva. Nel 1956, su iniziativa di John McCarthy, e insieme a Minsky e Shannon, diversi ricercatori si riunirono presso il Dartmouth College, negli Stati Uniti, per il primo seminario che di fatto segnò la nascita del campo dell'AI. Fu in tale occasione che si gettarono molte di quelle che in seguito sarebbero diventate le fondamenta classiche della materia.

IL MEDIOEVO DELLO SVILUPPO DELL'AI

Negli anni '60 il contributo più importante al settore fu probabilmente il General Problem Solver di Newell e Simon. Si trattava di un programma polivalente volto a simulare, tramite l'utilizzo di un computer, alcuni metodi umani di problem solving. Purtroppo la tecnica impiegata non era particolarmente efficace e, a causa del tempo impiegato e dei requisiti di memoria richiesti per risolvere problemi anche relativamente semplici, il progetto fu abbandonato.
L'altro contributo significativo di quegli anni fu quello di Lotfi Zadeh, che introdusse il concetto di insiemi e sistemi "fuzzy", concetto in base al quale i computer non devono operare secondo un formato logico esclusivamente binario, ma anche tramite processi "sfocati" che ricalcano quelli umani. Questa tecnica e i suoi derivati sono esaminati nel Capitolo 4.
Oltre a questi esempi, gli anni '60 diedero vita ad alcune temerarie rivendicazioni sul potenziale dell'AI nell'imitare e forse perfino ricreare l'intero funzionamento del cervello umano in un brevissimo lasso di tempo. Con il senno di

poi, va detto che cercare di far funzionare un computer nella stessa maniera di un cervello umano era un po' come cercare di far volare un aereo esattamente come un uccello. Nel secondo caso si rischierebbe di perdere quanto di buono c'è in un velivolo e, di fatto, in quel momento la ricerca sull'AI si lasciò sfuggire gran parte delle opportunità offerte dai computer.

Purtroppo (e abbastanza sorprendentemente), certe limitazioni di pensiero persistono ancora oggi. Alcuni degli attuali libri di testo (tanti anche sotto le etichette dell'AI moderna) continuano a focalizzarsi esclusivamente sull'approccio classico, quello mirato a far sì che un computer imiti l'intelligenza umana, senza tenere realmente in conto l'entità e le entusiasmanti possibilità dei diversi tipi di AI (considerate come macchine intelligenti di per sé, e non semplicemente in grado di replicare l'intelligenza umana).

In quel periodo, notevoli sforzi furono orientati a far comprendere e parlare ai computer il linguaggio umano naturale, piuttosto che, in maniera più diretta, il loro linguaggio macchina. Questa tendenza fu in parte ispirata dalla concezione dell'intelligenza proposta da Turing, ma anche dal desiderio che i computer si interfacciassero più facilmente con il mondo reale.

Uno dei migliori programmi per computer in lingua inglese, ELIZA, fu messo a punto da Joseph Weisenbaum. A dire il vero, fu addirittura il primo esemplare di quelli che sarebbero poi diventati noti come "chatterbots".[2]

Anche in questa fase relativamente precoce, alcune delle

[2] Da chat (chiacchierare) e bot (robot): software in grado di simulare una conversazione con esseri umani

conversazioni che intratteneva erano sufficientemente realistiche da indurre, di tanto in tanto, alcuni utenti a pensare che quello con cui stavano comunicando fosse un essere umano e non un computer.
In realtà, ELIZA forniva in genere risposte preconfezionate o, semplicemente, ripeteva ciò che gli era stato detto, riformulando la risposta con l'uso di alcune regole grammaticali di base. Tuttavia, anche questa operazione sembrava riprodurre adeguatamente, in qualche misura, alcune delle attività colloquiali tipiche degli esseri umani.

L'EPOCA BUIA DELLA RICERCA SULL'AI

Dopo il fermento degli anni '60, periodo di sostanziosi finanziamenti alla ricerca e di annunci sui traguardi che sarebbero stati raggiunti a breve in termini di capacità da parte dell'AI di replicare quella umana, il decennio successivo si rivelò deludente, configurandosi per molti versi come un'epoca buia per il settore. Alcune delle rivendicazioni più ottimistiche degli anni '60 avevano portato le aspettative a un livello estremo, e quando i risultati promessi non si materializzarono, gran parte dei finanziamenti per la ricerca sull'AI vennero meno.
Al tempo stesso, il campo delle reti neurali – ovvero dei computer che replicavano la struttura neurale del cervello – si arrestò quasi da un giorno all'altro per via di un feroce attacco di Marvin Minsky e Seymour Papert all'incapacità dei percettroni di generalizzare per affrontare determinati tipi di problemi relativamente semplici (argomento che affronteremo nel Capitolo 4).
Bisogna certo tener conto, però, che negli anni '70 le capacità dei computer e, di conseguenza, dei programmi AI erano piuttosto limitate rispetto a quelle attuali. Anche i

programmi migliori potevano affrontare solo le versioni semplici dei problemi che avevano lo scopo di risolvere; in effetti, a quel tempo tutti i programmi erano, in un certo senso, programmi "giocattolo".

I ricercatori si erano infatti scontrati con diversi limiti fondamentali che avrebbero superato solo molto più tardi. Il più importante di questi era la limitata potenza di calcolo.

Velocità e memoria disponibili non erano nemmeno lontanamente sufficienti per eseguire compiti di una qualche reale utilità: si pensi per esempio al linguaggio macchina naturale di Ross Quillan, che doveva cavarsela con un vocabolario di venti parole!

Il problema principale era però che i compiti dell'AI – come far sì che un computer comunicasse in un linguaggio naturale o comprendesse il contenuto di un'immagine in maniera simile agli esseri umani – richiedevano molte informazioni e notevole potenza di elaborazione, anche per operare a un livello molto basso e ristretto. Per un computer può essere difficile distinguere un'immagine (anche ritraente oggetti ordinari e di uso quotidiano), e ciò che gli umani considerano come un ragionamento di senso comune su parole e oggetti richiede in realtà un sacco di informazioni generiche.

Come se le difficoltà tecniche incontrate negli anni '70 non fossero abbastanza, il campo divenne anche argomento di interesse per i filosofi. John Searle, per esempio, propose il suo problema della stanza cinese (lo incontreremo nel Capitolo 3) per dimostrare come non sia possibile sostenere che un computer "comprenda" i simboli con cui comunica. Di conseguenza, la macchina non può necessariamente essere considerata "pensante" – come postulato in precedenza da Turing – solo perché è in grado di manipolare dei simboli.

Benché molti ricercatori pratici portassero semplicemente avanti il loro lavoro evitando le critiche, diversi filosofi (come, appunto, Searle) insistettero nel sostenere che i risultati effettivi dell'AI sarebbero stati sempre fortemente limitati. Riferendosi proprio a questi ultimi, Minsky dichiarò: "Fraintendono, ragion per cui dovrebbero essere ignorati". Fu dunque un periodo di aspri conflitti, che distolsero l'attenzione dagli sviluppi tecnici, per dirottarla verso diatribe filosofiche che (con il senno di poi) oggi molti considerano false piste.

All'epoca quasi isolato, John McCarthy riteneva invece che il modo in cui opera il cervello umano (e in cui agiscono gli umani) non avesse una diretta rilevanza per l'AI. Era cioè dell'idea che tutto ciò di cui c'era davvero bisogno fossero macchine in grado di risolvere i problemi, e non necessariamente computer capaci di pensare nello stesso identico modo in cui pensa la gente. Minsky criticò questo punto di vista, affermando che una corretta comprensione degli oggetti o una corretta conversazione richiedessero un computer capace di pensare come una persona. E così la disputa andò avanti…

LA RINASCITA DELL'AI

Gli anni '80 videro una sorta di rinascita dell'AI, e questo essenzialmente per tre fattori.

In primo luogo, molti ricercatori seguirono l'esempio di McCarthy e continuarono a perfezionare sistemi AI partendo da un punto di vista pratico. Si assistette perciò allo sviluppo di "sistemi esperti" progettati per affrontare ambiti molto specifici di conoscenza, evitando dunque in qualche in modo gli argomenti sulla base di una mancanza di "buon senso". Elaborati inizialmente negli anni '70, que-

sti sistemi trovarono le prime vere applicazioni pratiche nel settore industriale solo nel decennio successivo.
In secondo luogo, nonostante le discussioni (e le controversie) filosofiche proseguissero, in particolare sulle probabilità che una macchina potesse pensare allo stesso modo di un essere umano, lo fecero però nel modo più indipendente possibile dal percorso pratico che nel frattempo l'AI aveva intrapreso. Le due scuole procedettero semplicemente in parallelo, e gli sviluppatori si concentrarono sulle soluzioni concrete in campo industriale senza sostenere necessariamente che i computer dovessero o potessero comportarsi come gli esseri umani.
In terzo luogo, il concomitante sviluppo della robotica cominciò a esercitare una notevole influenza sull'AI. A tal proposito emerse un nuovo paradigma nella convinzione che, per mostrare un'intelligenza "reale", un computer dovesse avere un corpo in grado di percepire, muoversi nello spazio e sopravvivere nel mondo. Senza tali capacità, si sosteneva, come era lecito attendersi che un computer potesse comportarsi alla stessa maniera di un essere umano? O che sperimentasse il senso comune? Il nuovo influsso della cibernetica, dunque, mise molta più enfasi nella realizzazione di AI tramite un approccio bottom-up; lo stesso, in effetti, originariamente ipotizzato da McCulloch e Pitts.

L'AI AI GIORNI NOSTRI

Poco alla volta, il campo emergente dell'AI cominciò ad acquisire sicurezza. Le sue applicazioni industriali crebbero di numero e iniziarono a essere utilizzate in aree sempre più ampie, come i sistemi finanziari e militari. Aree in cui dimostrò di poter sostituire un operatore umano, ma anche,

in molti casi, di offrire prestazioni di gran lunga migliori. Le applicazioni dell'AI in queste aree si sono ormai estese enormemente, al punto che società finanziarie che prima facevano soldi grazie alle consulenze ai clienti ricavano adesso profitti molto più lauti sviluppando sistemi AI che vendono o mettono al servizio della clientela.

A partire dai primi anni '90 si è assistito anche al raggiungimento di diverse tappe importanti. L'11 maggio 1997, per esempio, Deep Blue è stato il primo computer a battere il campione del mondo in carica di scacchi (Garry Kasparov). In un altro ambito, il 14 marzo 2002 Kevin Warwick (l'autore di questo libro) è stato il primo a collegare con successo il sistema nervoso umano direttamente a un computer per dare istantaneamente vita a una nuova forma combinata di AI, se non a qualcosa di più. L'8 ottobre 2005 è arrivato l'annuncio che un robot della Stanford University aveva vinto il DARPA Grand Challenge guidando in autonomia per 131 miglia lungo un percorso improvvisato nel deserto. Nel 2009, infine, il team del progetto Blue Brain ha dato notizia di aver simulato con successo alcune aree della corteccia di un ratto.

Nella maggior parte dei casi, simili affermazioni non derivano da tecnologie di recente invenzione, quanto piuttosto dall'aver spostato i limiti della tecnologia disponibile. In effetti, Deep Blue era oltre dieci milioni di volte più veloce rispetto al primo computer Ferranti (1951) capace di giocare a scacchi. Lo straordinario e ininterrotto incremento, anno dopo anno, della potenza di calcolo si è attenuto a quella che è ormai nota come legge di Moore, e che lo ha al tempo stesso previsto.

Secondo la legge di Moore, la velocità e la capacità di memoria dei computer raddoppia ogni due anni. Ciò significa che

i problemi incontrati dai precedenti sistemi AI vengono superati abbastanza rapidamente dalla potenza di calcolo pura e semplice. È interessante notare come, anno dopo anno, c'è sempre qualcuno pronto a sostenere su qualche giornale che la legge di Moore sarà superata da fattori limitanti come la dimensione, il calore, il costo, eccetera. A ogni modo, i continui progressi tecnologici fanno sì che ogni anno la potenza di calcolo disponibile raddoppi e che la legge di Moore continui a rivelarsi valida.

Inoltre, questi ultimi anni hanno visto l'emergere di approcci insoliti all'AI. Un esempio è il metodo degli "agenti intelligenti". Si tratta di un approccio modulare che, in qualche modo, sembra imitare il cervello mettendo insieme diversi agenti specializzati per affrontare ciascun problema, alla stessa maniera in cui un cervello dispone di diverse aree da utilizzare in situazioni differenti. Questo approccio si adatta perfettamente anche ai metodi informatici in cui diversi programmi vengono associati – come richiesto – a diversi oggetti o moduli.

Un agente intelligente è molto più di un semplice programma. È un sistema di per sé, in quanto deve percepire il proprio ambiente e agire per massimizzare le possibilità di successo. Detto questo, è vero che, nella loro forma più elementare, gli agenti intelligenti sono solo programmi che risolvono problemi specifici. Tuttavia, questi agenti possono essere singoli sistemi robotizzati o macchine capaci di operare in maniera fisicamente autonoma.

Come descritto nel Capitolo 4, agenti a parte, in questo periodo sono emersi molti altri nuovi approcci nel campo dell'AI. Alcuni di questi sono stati di natura decisamente più matematica, come la probabilità e la teoria delle decisioni. Nel frattempo, il ruolo delle reti neurali e dei concetti

legati all'evoluzione, come gli algoritmi genetici, è diventato molto più influente.
È certamente il caso di particolari azioni che possono essere interpretate come intelligenti (sia negli umani, sia negli animali) e che possono essere eseguite (spesso in maniera più efficace) da un computer. È anche il caso di molti nuovi sviluppi dell'AI che hanno trovato applicazioni più generali, perdendo spesso, così, l'etichetta di "AI". Validi esempi di quanto appena detto sono il *data mining*,[3] il riconoscimento vocale e larga parte dell'attività decisionale attualmente svolta nel settore bancario. In ogni caso, ciò che in origine era AI si considera ormai semplicemente come parte di un programma informatico.

L'AVVENTO DEL WIRELESS

Una delle tecnologie chiave divenute realtà negli anni '90, a seguito della diffusione di Internet e del suo uso su larga scala, è stata il wireless come forma di comunicazione per i computer. Dal punto di vista dell'AI, questo ha cambiato completamente le regole del gioco. Fino a quel momento c'erano solo computer stand-alone, la cui potenza e le cui capacità potevano essere confrontate direttamente con i singoli cervelli umani (quella che è la loro "normale" configurazione). Nel momento in cui i computer in rete si sono affermati come la norma, piuttosto che considerare ciascuno di essi separatamente è diventato realisticamente necessario prendere in considerazione l'intera rete come un solo, grande cervello intelligente con ampia distribuzione (per questo detta appunto "intelligenza distribuita").

[3] Ricerca ed estrazione dei dati

Grazie alla tecnologia wireless, la connettività garantisce un enorme vantaggio per l'AI sull'intelligenza umana, nella sua attuale forma "stand-alone". All'inizio il wireless era principalmente un mezzo tramite il quale i computer potevano comunicare velocemente tra loro. Nondimeno, ha finito per disperdere in rete larghe sacche di memoria, facendo sì che la specializzazione si diffondesse e che l'informazione fluisse liberamente e rapidamente. Questo ha cambiato la prospettiva umana su sicurezza e privacy, e ha modificato le gerarchie tra i mezzi di comunicazione utilizzati dagli umani.

HAL 9000

Nel 1968 Arthur C. Clarke scrisse *2001: Odissea nello spazio*, da cui qualche tempo dopo Stanley Kubrick avrebbe tratto l'omonimo film. Uno dei protagonisti del film è HAL 9000, una macchina dotata di intelligenza pari o superiore a quella degli esseri umani. In effetti, sfoggia emozioni e capacità filosofiche tipicamente umane. Pur essendo niente più che un congegno immaginario, HAL divenne una sorta di traguardo a cui aspirare nel campo dell'AI. Alla fine degli anni '60 in parecchi credevano che una simile forma di AI sarebbe stata una realtà entro il 2001, soprattutto in virtù del fatto che HAL si basava sui dati scientifici del tempo.

In molti si sono chiesti come mai nel 2001 non siamo arrivati a una qualche forma di HAL, o almeno a una sua buona approssimazione. Minsky si è lamentato del troppo tempo perso dietro l'informatica industriale e sottratto alla comprensione di base di questioni come il senso comune. Analogamente, altri hanno deplorato che la ricerca sull'AI si concentrasse su semplici modelli neuronali, come il percettrone, piuttosto che sul tentativo di arrivare a un modello molto più vicino alle cellule originali del cervello umano.

Forse la risposta al perché non siamo arrivati a HAL entro il 2001 è un amalgama di tutti questi problemi e di altri ancora. Molto semplicemente, ci è mancato un impulso focalizzato a raggiungere questo tipo di AI. Nessuno ha investito i soldi per farlo e nessun team di ricerca ha lavorato al progetto. Sotto determinati aspetti – come il networking, la memoria e la velocità – avevamo già realizzato qualcosa di molto più potente di HAL prima del 2001, ma i riflessi emozionali e umorali in un computer non avevano (e probabilmente non hanno ancora) un ruolo peculiare da svolgere, se non forse nei film.

Per il guru Ray Kurzweil, il motivo del mancato avvento di HAL è questione di pura e semplice potenza computazionale e, in base alla legge di Moore, la sua previsione è che macchine con intelligenza di livello umano faranno la loro comparsa entro il 2029.

Naturalmente, cosa si intenda per "intelligenza di livello umano" rimane una domanda senza risposta. Io stesso, in un libro precedente, *March of the Machines*, ho azzardato una previsione non troppo distante da quella di Kurzweil, sostenendo che entro il 2050 le macchine disporranno di un'intelligenza che gli umani non saranno più in grado di gestire.

VERSO IL FUTURO

Gran parte della filosofia classica dell'AI (come vedremo nel Capitolo 3) si basa principalmente sul concetto di cervello o computer in quanto entità a sé stanti: un cervello disincarnato e posto in un barattolo, per così dire. Nel mondo reale, però, gli esseri umani interagiscono con il mondo che li circonda tramite gli organi sensoriali e le capacità motorie.

Ciò che oggi riveste un interesse notevole, destinato a cre-

scere in futuro, è l'effetto di un determinato corpo sulle capacità intellettuali del cervello di quel corpo stesso. La ricerca in corso mira a realizzare un sistema AI in un corpo – **EMBODIMENT**[4] – perché questo possa sperimentare il mondo, che si tratti della versione reale o di un mondo virtuale o addirittura simulato. Sebbene lo studio sia ancora focalizzato sul cervello AI in questione, il fatto che questo abbia un corpo con cui interagire con il mondo è un aspetto importante.

Proiettandoci verso il futuro, forse l'area più elettrizzante della ricerca sull'AI è quella in cui i cervelli AI si ottengono dal tessuto neurale biologico, in genere di un topo o di un essere umano. Le procedure previste e le modalità necessarie per riciclare e ottenere con successo tessuto neurale biologico vivo sono riportate in dettaglio nel Capitolo 5. In questo caso, l'AI non è più basata su un sistema informatico, quanto piuttosto su un cervello biologico riprodotto ciclicamente.

Il tema è di sicuro interesse, trattandosi di una nuova forma di AI, e sarà potenzialmente utile, in futuro, per i robot domestici. Tuttavia, dà il "la" anche a una nuova, significativa, area di studi, dal momento che mette in discussione molti dei presupposti filosofici su cui si fonda l'AI classica. Ciò che mette in discussione, in sostanza, è la differenza tra l'intelligenza umana e quella di una macchina di silicio. In questo nuovo settore di ricerca, però, i cervelli AI possono essere ricavati a partire da neuroni umani e sviluppati in qualcosa di simile a una versione AI del cervello umano,

[4] Processo mediante cui si dota un cervello o una rete neurale artificiale di un corpo fisico così che possa interagire con il mondo reale

sfumando così i contorni di quello che era un divario netto tra due modelli decisamente differenti di cervello.

CYBORG

Si potrebbe sostenere che, quando a un cervello AI biologico viene assegnato un corpo robotico, quello che si ottiene è una specie di cyborg – un organismo cibernetico (parte animale/umano, parte tecnologia/macchina) – con un cervello "embodied".[5] Quest'area di ricerca è la più elettrizzante di tutte: la diretta connessione tra un animale e una macchina per il miglioramento (in termini prestazionali) di entrambi. Un simile cyborg è solo una versione potenziale. Infatti, né il cyborg su cui si concentra attualmente la ricerca, né quello solitamente incontrato nelle opere di fantascienza è di questo tipo.

Il genere di cyborg in cui ci si imbatte con maggior regolarità è quello in forma di essere umano con impiantata una tecnologia integrata e collegata a un computer che gli fornisce, così, capacità superiori alla norma umana (in altre parole: un cyborg possiede capacità che un essere umano non ha). Queste capacità possono essere fisiche e/o mentali e possono avere a che fare con l'intelligenza. In particolare, vedremo che un cervello AI è solitamente (tranne nel caso in cui si tratti di cervelli biologici AI) molto diverso da un cervello umano, e queste differenze possono realizzarsi in termini di vantaggi (soprattutto per l'AI).

I motivi sottesi alla creazione di cyborg ruotano generalmente intorno al tentativo di potenziare le prestazioni del cervello umano collegandolo direttamente a una macchina.

[5] Vedi nota 4

Il cervello combinato può quindi, almeno potenzialmente, funzionare con caratteristiche di entrambe le sue componenti: un cyborg potrebbe dunque avere una memoria migliore, maggiori capacità matematiche, sensi più acuti, un pensiero multidimensionale e migliori capacità di comunicazione rispetto a un cervello umano. Fino a oggi, gli esperimenti hanno dato risultati positivi sia per quanto riguarda il potenziamento sensoriale, sia per ciò che concerne una nuova forma di comunicazione per cyborg. Sebbene l'argomento non sia specificatamente trattato in questa sede, ritengo che il materiale esposto nei Capitoli 5 e 6 metterà il lettore in una buona posizione per proseguire gli studi in questo settore.

IN CONCLUSIONE

Questo capitolo ha preparato il terreno per il resto del libro, illustrando una breve panoramica sull'evoluzione storica dell'AI e su alcuni dei principali sviluppi, e introducendo, in tal modo, alcune delle personalità influenti del settore.
Nel prosieguo dell'opera, dopo una semplice introduzione (Capitolo 1) al concetto di intelligenza nel suo complesso, i Capitoli 2 e 3 si concentreranno sui metodi – più datati – di AI classica. I Capitoli 4 e 5 prenderanno invece in considerazione gli approcci moderni, attuali e più futuristici. Avrete modo di vedere come il materiale trattato nelle sezioni più aggiornate dei Capitoli 4 e 5 non compaia probabilmente nella maggior parte degli altri manuali sull'AI (anche quando i suddetti manuali sfoggiano titoli come *Intelligenza Artificiale* o *AI: Un approccio moderno*). Il Capitolo 6 indaga infine su come una AI sia in grado di percepire il mondo attraverso il suo sistema di sensori.
Buon divertimento!