

ERNESTO D'AVANZO

DEMONI LAPLACIANI, MENTI
«INCARNATE» E «DISINCARNATE»
«CASSETTE DEGLI ATTREZZI»
PER PROBLEMI FILOSOFICI

ABSTRACT - D'AVANZO E., 2016 - Laplace's demons, embodied and disembodied minds. Toolboxes for philosophical problems.

Atti Acc. Rov. Agiati, a. 266, 2016, ser. IX, vol. VI, B: 65-78.

Plato proposed the dualistic solution to the *mind-body* problem, providing an explanation along the lines of his epistemology. While Francis Bacon, in 1600, formulated his vision of the *scientific method* that will be valid until the 60's of last century, when Karl Popper proposed his version, entering into controversy with the Lord Chancellor. The *mind-body* problem and the *scientific method* are just two of the issues concerning the foundations of the sciences that have been addressed, mainly, with conceptual tools. Thanks to recent developments in artificial intelligence and computational neuroscience these problems have new empirical tools to be analyzed. The present discussion is an attempt to tell the birth, development, and some examples of these toolboxes, available to all those who want to apply knowledge to improve society.

KEY WORDS - Laplace demon; Mind-body problem; Neuroeconomy; Artificial intelligence; Functionalism.

RIASSUNTO - D'AVANZO E., 2016 - Demoni laplaciani, menti «incarnate» e «disincarnate». «Cassette degli attrezzi» per problemi filosofici.

Platone propose la soluzione *dualistica* al problema *mente-corpo*, fornendo una spiegazione alla luce della sua *epistemologia*. Francesco Bacone, invece, nel 1600 formulò la sua visione del metodo scientifico che sarà valida fino agli anni '60 del secolo scorso, quando Karl Popper propone la sua versione, entrando in polemica con il Lord cancelliere. Il problema mente-corpo e il metodo scientifico sono solo due dei temi riguardanti i fondamenti delle scienze che sono stati affrontati, prevalentemente, con strumenti concettuali. Grazie ai recenti sviluppi dell'intelligenza artificiale e delle neuroscienze computazionali molti problemi dispongono di nuovi strumenti, questa volta empirici, per analizzare temi sui fondamenti della scienze. La presente trattazione è un tentativo, non esaustivo, per raccontare la nascita, lo sviluppo e alcuni esempi di queste nuove «cassette degli attrezzi», a disposizione di tutti coloro che hanno voglia di applicare la conoscenza per migliorare la società.

PAROLE CHIAVE - Demone di Laplace; Problema mente-corpo; Neuroeconomia; Intelligenza artificiale; Funzionalismo.

PROLOGO

Qual è la relazione esistente fra gli eventi mentali e gli eventi fisici? La domanda identifica il *problema mente-corpo* tanto dibattuto in filosofia della mente, e in tante discipline scientifiche, come la psicologia, la fisiologia, e più recentemente, le neuroscienze e la scienza cognitiva. Platone propose la soluzione *dualistica* al problema mente-corpo: mente e corpo, o mente e cervello, sono due sostanze separate e incapaci di interagire l'una con l'altra. Il filosofo greco, naturalmente, fornì una spiegazione del problema sulla base della sua *epistemologia*, identificando la mente con l'anima e sostenendo che la mente fosse pre-esistente al corpo e che, dopo la morte, l'anima sopravvivesse.

Eppure, le due entità postulate da Platone, anima e corpo, o se si preferisce, mente e cervello, interagiscono, questo è evidente.

Nonostante ciò, la soluzione platonica al problema non spiega come la volontà muove il corpo, in altri termini, come un'entità interagisce con l'altra.

La filosofia, da Platone sino a oggi, ha provato a dare spiegazioni del problema con strumenti concettuali. La scienza, anch'essa impegnata su tale fronte, solo dal 1600 ha cominciato a fornire spiegazioni del problema mente-cervello in modo più sistematico, e facendo ricorso a strumenti empirici e/o formali. Cartesio, ad esempio, riconosce l'esistenza di due sostanze, *mente e corpo*, note anche come «*res cogitans*» e «*res extensa*», e introduce l'elemento *meccanicistico*, ossia l'*azione riflessa* che, nella sua versione moderna, è un sistema di controllo biologico il quale collega stimolo e risposta mediante l'arco riflesso. Il sistema contempla un organo recettore, uno effettore e un canale di comunicazione fra i due.

Un altro tentativo sistematico che affronta il problema risale agli inizi del 1900, quando, Gottlob Frege, Bertrand Russell e David Hilbert, nel cercare di sistematizzare la matematica su base logica, cercarono di descrivere, in tali termini, ogni sistema fisico, compreso il cervello. Modelli del sistema nervoso potevano trovare nella logica matematica uno strumento per capire come i segnali provenienti dagli organi recettori, gli stimoli, ad esempio la «volontà» o l'ordine di andare a fare la spesa, causino le risposte degli organi effettori, ad esempio muovere le gambe, e così via.

Quanto detto sopra rende l'idea di come il problema mente-corpo, da sempre, abbia suscitato l'interesse delle menti più brillanti. Gli studiosi che se ne sono interessati hanno visto in esso una sfida per i metodi empirici e/o matematici proposti per altri settori della scienza, come il marchese di Laplace, più noto per le sue indagini in campo fisico/astronomico ma che ha posto, proprio nelle sue teorie scientifiche, il problema di un *intelletto*

dalle immense capacità, e completamente deterministico, tanto da meritare l'appellativo di *demone*. Il demone di Laplace è introdotto nella Sezione 2, contestualmente ad alcuni accenni alle due principali opere in campo fisico del marchese, e alle implicazioni che esse hanno per il problema mente-corpo e il *libero arbitrio*, due argomenti strettamente legati. La Sezione 3 analizza cosa accade quando si passa da una posizione dualistica del problema, à la Platone, alla posizione monistica, che vede la mente identificata, o "incarnata", con il cervello. Uno sguardo a un programma di ricerca empirico, quello proposto dalle *neuroscienze*, con alcuni riferimenti alle sue applicazioni nel campo delle scelte economiche, offre una visione differente del problema mente-corpo, almeno rispetto a quella che viene fuori dalla visione di Laplace. Nella Sezione 4, si accenna, brevemente, all'*intelligenza artificiale* e alla corrispondente *teoria della mente*, programma della scienza cognitiva. La Sezione 5 presenta un breve quadro di due programmi per calcolatore, *Remote agent* e *BACON* che rappresentano i prototipi di alcune categorie di programmi successivi, e di successo, i quali fanno ampio uso di metodi di apprendimento automatico e, in generale, di intelligenza artificiale. L'intelligenza artificiale si è rivelato uno strumento utilissimo per affrontare altri temi riguardanti i fondamenti delle scienze, come il «metodo scientifico», presentato brevemente nella Sezione 6.

Come ogni cammino «catartico» il racconto comincia dagli inferi e dai suoi abitanti, i demoni.

IL DEMONE DI LAPLACE: UN INTELLETTO ONNISCIENTE

Allen Downey, nel suo *Think Complexity* (2012), riprende un celebre passaggio scritto da *Pierre-Simon* marchese di *Laplace*, nel quale lo scienziato francese, ispirato dalle sue scoperte in analisi matematica e in astronomia, sostiene che «lo stato attuale dell'universo può essere considerato come l'*effetto* del suo passato e la *causa* del suo futuro». Per Laplace un «intelletto» a conoscenza di «tutte le forze che mettono in moto la natura», «delle posizioni di tutti gli oggetti di cui la natura è composta» ebbene, per un tale intelletto, «ampio a sufficienza da potere analizzare tutti questi dati», «nulla sarebbe incerto e il futuro, proprio come il passato, sarebbe evidente davanti ai suoi occhi».

Il brano citato da Downey, è tratto dall'*Essai philosophique sur les probabilit *, un'opera del 1814, che Laplace, allora sessantacinquenne, cura contemporaneamente alla stesura della pi  impegnativa, e controversa, *M canique C leste*, iniziata nel 1799 e terminata un paio d'anni prima della sua morte, avvenuta nel marzo del 1827.

Sebbene la *Mécanique* susciti interesse soprattutto per la sua valenza nel descrivere l'impresa della fisica astronomica, iniziata con Isaac Newton, nei termini rigorosi del calcolo differenziale, l'opera è, altresì, ricca di intuizioni riguardanti fenomeni che tutt'oggi destano l'interesse della comunità scientifica, non solo da un punto di vista sperimentale; basti pensare all'*ipotesi* sull'esistenza dei *buchi neri*; oppure all'*ipotesi della nebulosa*, secondo cui il sistema solare si sarebbe sviluppato a partire da una massa globulare di gas. Entrambi questi postulati sono il prodotto dello scienziato francese; l'*ipotesi della nebulosa*, in realtà, era già stata introdotta da Immanuel Kant nel 1755, ma il marchese di Laplace, qualche anno dopo, la riformula autonomamente rispetto al filosofo tedesco.

La lettura d'insieme, della *Mécanique* e dell'*Essai*, non è dettata solo dalla comune finestra temporale in cui sono prodotte le due opere; la comparazione ha una motivazione più profonda. Anzitutto va notato che uno dei maggiori contributi della *Mécanique*, come si accennava pocanzi, risiede proprio nello sviluppo di metodi matematici adatti alla formulazione delle teorie fisiche. Grazie a questi metodi, che oggi costituiscono la branca della *fisica matematica*, Laplace vede l'universo fisico a lui conosciuto, quello cosmologico, come un grande meccanismo di orologio, perfettamente *deterministico*, in cui, appunto, ogni "stato presente e futuro" è perfettamente determinato grazie alla conoscenza delle condizioni iniziali assunte dagli oggetti, e alle leggi che li governano, le quali sono scritte nel linguaggio rigoroso dell'analisi matematica. Il resto del lavoro sul *determinismo* viene fuori dalla stesura dell'*Essai*, un trattato il cui tema principale è la *teoria delle probabilità* ma che contiene un altro tema, anche esso di importanza fondamentale, e cioè il *ragionamento per induzione* per la cui esemplificazione i filosofi, talvolta, ricorrono all'ornitologia: *ho visto un corvo nero; ho visto un secondo corvo nero; e così via*. La conclusione per *induzione* è che *il prossimo corvo sarà nero*. Nella versione probabilistica dell'*induzione*, il prossimo corvo *probabilmente* sarà nero.

L'applicazione del *ragionamento per induzione* alla formulazione dei fenomeni cosmologici, fatta da Laplace, ed espressa in termini di equazioni differenziali, costituisce il nucleo del suo *determinismo*, richiamato nella citazione fatta all'inizio. Nel caso di Laplace, al posto dei corvi usati dai filosofi, ci sono gli oggetti del cosmo e le equazioni, le cui soluzioni rappresentano i moti degli oggetti stessi.

Come si diceva sopra, uno degli aspetti più dibattuti e controversi del *determinismo* di Laplace è rappresentato dal suo «intelletto» il quale, se conoscesse «tutte le forze che mettono in moto la natura», e le «posizioni di tutti gli oggetti di cui la natura è composta», avrebbe «davanti ai suoi occhi» il futuro come il passato.

Se Laplace abbia mai creduto, effettivamente, all'esistenza di un tale *intelletto*, cui biografi e critici successivi si riferiranno come al *demone di Laplace*, è materia di discussione. Giacché il marchese era dichiaratamente agnostico, la sua analisi, molto probabilmente, non riguardava l'*onniscienza* come qualità dell'intelletto di Dio. Nel caso, invece, che la sua attenzione fosse rivolta alle capacità umane, o alla possibilità che tali calcoli potessero essere eseguiti da un *intelletto artificiale*, come un programma per calcolatore, ebbene, alcune teorie fisiche del XIX e XX secolo, quali l'*entropia*, il *decadimento radioattivo* o la *meccanica quantistica*, hanno abbondantemente smentito una tale ipotesi, lanciando, al tempo stesso, nuove sfide al *determinismo*.

Una visione deterministica della realtà rappresenta, altresì, una sfida nei confronti dell'abilità umana di fare scelte e prendere decisioni o, più sinteticamente, dell'esistenza, o meno, del *libero arbitrio*. Infatti, se il nostro corpo fisico, compreso il nostro cervello, è governato da leggi fisiche, come può permettere il *libero arbitrio*? La risposta a tale domanda, ancora oggi non definitiva, ha visto l'impiego di notevoli risorse intellettuali e sperimentali, come dimostrano i racconti delle sezioni seguenti.

MENTI +O- INCARNATE E LA NEUROECONOMIA

Il maggior contributo di Charles Sherrington (1910) consiste nell'aver dato alla fisiologia la possibilità di associare processi cerebrali ai comportamenti umani: per ogni attività mentale esiste un preciso processo cerebrale sotteso. Il sistema nervoso concepito dal neurofisiologo inglese, tuttavia, non spiega come l'attività mentale, e le idee, ad esempio la volontà, possa causare il movimento del corpo.

Weiss e Bernstein (2014), fra i più autorevoli critici della teoria di Sherrington, sostennero l'idea che il comportamento umano è organizzato secondo «obiettivi» e non attivazioni muscolari come sostenuto da Sherrington e, prima di lui, da Cartesio. Per comprendere la fisiologia del sistema nervoso, quindi, sarebbe stato necessario capire le computazioni che il sistema nervoso esegue.

In un articolo apparso nel 1976, e divenuto popolarissimo, David Marr e Tomaso Poggio, invece di isolare un ristretto comportamento definendo poi, in termini sherringtoniani, l'arco riflesso che lo ha prodotto, descrivono cosa fa il sistema nel suo insieme, per poi studiare, in termini matematici, l'hardware biologico che ha raggiunto quell'obiettivo. Quando si studia la relazione fra cervello e comportamento, sostiene Marr, è la funzione del sistema nervoso che deve essere oggetto della ricerca. Il contributo di Marr alla

scienza *neurocomputazionale* è di aver introdotto la descrizione del sistema nervoso in conformità a obiettivi da seguire, sebbene il neuroscienziato non abbia prestato tantissima attenzione a sviluppare un apparato sperimentale per specificare tali obiettivi. Tomaso Poggio, uno dei collaboratori di Marr, invece, ha applicato, sistematicamente, l'approccio proposto da Marr, proponendo un modello computazionale della corteccia visiva.

L'idea è basata sulla convinzione che le abilità delle nostre facoltà visive sono evidenti. Da tale semplice assunto, Poggio realizza una struttura computazionale del processo neurobiologico che permetta a una macchina, su cui il modello è cablato, di avere abilità simili a quelle esibite da un essere umano per lo stesso compito. Poggio è uno degli allievi di Marr che, fra i primi, realizza l'*occhio artificiale* simulato via software su un computer, in grado di riconoscere immagini anche in movimento. Il modello proposto da Marr e Poggio, e tutte le successive varianti, non solo ha trovato una vastità di applicazioni pratiche, come nella robotica industriale, la guida autonoma dei veicoli o l'analisi delle immagini mediche, solo per citare alcuni esempi, ma ha anche una sua specificità metodologica che rappresenta una risposta per tutti coloro i quali, da un punto di vista epistemologico, hanno considerato l'intelligenza artificiale, e i campi di indagine a essa correlati, come una «pratica ingegneristica». Il modello proposto, infatti, è ispirato biologicamente, cioè Poggio l'ha ottenuto osservando i fenomeni che avvengono all'interno dell'area visiva del cervello. Ha, per così dire, indagato le funzioni visive nel loro luogo "naturale".

Seguendo il suggerimento di Marr molti studiosi sostengono che per analizzare la «funzione del sistema nervoso» è sufficiente guardare al processo con cui prendiamo decisioni, la cui analisi, nel caso del singolo individuo, è compito della *teoria delle decisioni*, mentre, laddove entrino in gioco più individui, rientra nella più ampia *teoria dei giochi*. Un aspetto interessante di questo programma di ricerca, meglio conosciuto come neuroeconomia, è l'uso che esso fa dello strumento del calcolo delle probabilità per trattare le cosiddette decisioni in situazioni di incertezza. Partendo dagli obiettivi e dalle decisioni che singoli o gruppi di individui prendono è possibile risalire ai fenomeni neurobiologici che sottendono queste decisioni, mediante un'analisi dall'alto verso il basso, facendo uso di tecniche sperimentali di laboratorio come il neuroimaging (TAC, PET o risonanze magnetiche funzionali). Per identificare un obiettivo, quindi, prima bisogna essere capaci di calcolare l'utilità di ciascuna azione, valutandone poi la probabilità secondo alcuni parametri. Come in ogni programma scientifico sarà «l'evidenza empirica» che esso riuscirà a fornire a legittimare la *neuroeconomia*, come sostiene Paul Glimcher, uno dei maggiori esponenti della neonata disciplina. Naturalmente i ricercatori e gli studiosi coinvolti devono dare anche tante

spiegazioni teoriche, come alcuni assunti mutuati dall'ecologia evolutiva.

Oltre agli aspetti meramente scientifici, il programma neuroeconomico introduce nel dibattito, non solo filosofico, appassionanti discussioni riguardanti il libero arbitrio. Introducendo la matematica probabilistica come metodo d'indagine del rapporto mente-cervello, la neuroeconomia offre una visione completamente differente del *libero arbitrio*. Ciò che è messo in discussione è la prevedibilità e la determinatezza delle nostre decisioni; gli stati futuri del mondo e degli organismi che lo abitano, comprese le decisioni degli esseri umani, non sono pre-determinati. Da questo punto di vista il mondo è un sistema indeterminato dove gli uomini sono liberi di scegliere.

L'IA E LE MENTI DISINCARNATE

Il 1956, anno in cui si è tenuta la scuola estiva a Dartmouth, negli Stati Uniti, è unanimemente riconosciuto come la data di nascita ufficiale dell'*intelligenza artificiale* (IA). Allen Newell, Clifford Shaw, Herbert Simon, Marvin Minsky e altri studiosi, di diversa provenienza scientifica, iniziarono, proprio a Dartmouth, la realizzazione di «macchine», ovvero programmi, oggi diremmo software, capaci di simulare comportamenti che in qualche modo potessero essere considerati «intelligenti». L'obiettivo dei padri fondatori era «costruire artefatti con prestazioni tali da assistere l'uomo nei compiti teorici o pratici di diversa complessità».

Tra i «padri fondatori» era diffuso un criterio per verificare l'intelligenza delle macchine che si realizzavano, il *test di Turing*, noto anche come *imitation game*, dal nome del suo inventore, il celebre matematico britannico Alan Turing (1950). Ecco una versione molto semplificata del *test*. Mettiamo un computer e un essere umano dietro un muro lasciandoli dialogare con un'altra persona, oltre il muro, che chiamiamo, per comodità, ispettore. All'ispettore è nota la presenza della macchina e dell'umano senza conoscere, tuttavia, la loro disposizione. Tramite una telescrivente l'ispettore può interrogare la macchina e l'essere umano, e se dalle risposte non riesce a distinguere l'una dall'altro, allora la macchina ha superato il *test* dell'intelligenza.

In realtà l'IA, oltre a realizzare artefatti «intelligenti», nella versione di Turing, si è posta, sin dall'inizio, il duplice e arduo scopo di fornire una spiegazione del funzionamento mentale, proponendo, altresì, una propria *teoria della mente*. E quando, per tale scopo, l'IA si è confrontata con la *psicologia*, la *linguistica*, la *filosofia della mente*, la *scienza dei calcolatori*, la *logica* e la *statistica*, ha contribuito a dare vita, non solo nel mondo accademico, alla

scienza cognitiva. Volendo riassumere in una formula molto semplificata, e controversa, per la *scienza cognitiva* la mente sta al cervello come per la *scienza dei calcolatori* il software (il programma) sta al suo hardware (la macchina fisica); in altre parole la mente è considerata un sistema per l'elaborazione dell'informazione. Così la neonata *scienza*, a ispirazione computazionale, ha dato da un lato spiegazioni di fenomeni mentali soprattutto riguardanti le capacità cognitive "alte", come, ad esempio, la soluzione di problemi, il ragionamento, la pianificazione, l'analisi del linguaggio naturale, la visione, e dall'altro ha introdotto una plausibilità cognitiva nella realizzazione di macchine che potessero simulare alcuni compiti umani.

Il contributo che l'IA ha dato alla realizzazione di macchine industriali di successo, tuttavia, ha indotto alcuni studiosi, come il filosofo John Searle, a considerare la disciplina più una "pratica sperimentale tra l'informatica e l'ingegneria" che una scienza. A determinare questo tipo di considerazione, tuttavia, sembra abbia giocato un ruolo fondamentale l'atteggiamento della *scienza cognitiva* nei confronti del cervello. Gli scienziati cognitivi, infatti, hanno "prediletto" lo studio dei *processi mentali* a scapito di quelli *cerebrali*, neurobiologici, rivolgendo, così, la propria indagine verso la mente "disincarnata".

Perché tale "predilezione"? Certamente non è stata una dimenticanza degli scienziati cognitivi. Nemmeno si tratta di una scelta di principio. È stata piuttosto una scelta metodologica obbligata e dettata dalla complessità dell'oggetto di studio: il cervello. Basta considerare il numero di cellule, i neuroni, che lo costituiscono, stimato nell'ordine di cento miliardi. Senza considerare, poi, che ciascun neurone ha diverse migliaia di collegamenti con gli altri neuroni. Problemi di tale complessità costituiscono seri ostacoli per i ricercatori, tanto che è stato coniato il termine anglosassone *curse of dimensionality*, la *maledizione della dimensionalità*, per descrivere i problemi che contano un numero altissimo di variabili in gioco, rendendo quindi molto difficile la ricerca della soluzione. In altri termini, per la *scienza cognitiva* l'apparato neurobiologico ha rappresentato sempre la sua maledizione, la *maledizione della neuroscienza*, dove per neuroscienza si intende lo studio del cervello a tutto tondo (neurobiologia, biochimica del cervello, elettrofisiologia, e così via). Sia chiaro che questo vale per la sola *scienza cognitiva* come programma di ricerca. Nel frattempo, infatti, sono stati proficui gli studi neuroscientifici, che hanno cominciato a far luce su tanti fenomeni e meccanismi cerebrali, come nel caso della visione artificiale, introdotto nella sezione precedente. La "maledizione", si diceva, è valsa soprattutto per l'IA, che ha sempre cercato di non considerare il cervello, al fine di ridurre la complessità della sua indagine, ma prestando attenzione, invece, agli "obiettivi", a cui si è accennato sopra, che hanno

costituito materia d'interesse per discipline quali la *teoria delle decisioni*, la *teoria dei giochi* o la *neuroeconomia*.

Per quanto riguarda il problema mente-corpo, è evidente che l'approccio dell'IA non è né di tipo *monistico* né di tipo *dualistico*: l'intelletto viene indagato su entrambi i livelli, fisico e mentale, ma privilegiando un approccio dall'alto in basso, *top-down*, indagando prima le funzioni dei processi, per poi esplorare i correlati neurobiologici. Non mancano le implicazioni per il *libero arbitrio* che, ricordiamo, è visto come l'abilità di creare opzioni e di scegliere fra di esse senza l'intervento di un agente esterno. Un programma per calcolatore, comunque complesso, che impieghi metodi di IA, come il *machine learning*, i quali gli permettano di cambiare uno o più dei suoi stati interni, autonomamente, cioè senza risorse esterne, ebbene un tale programma esibirebbe atteggiamenti ispirati da *libero arbitrio*.

Di seguito si introducono, seppur brevemente, alcuni casi studio provenienti dall'IA i quali rappresentano gli antesignani di tante applicazioni della nostra vita quotidiana che mostrano il successo di una disciplina da considerare tutt'altro che una "pratica ingegneristica". I casi presentati mostrano, ad esempio, come altri problemi classici riguardanti la scienza, come il dibattito sul *metodo scientifico*, hanno trovato nuovi percorsi esplorativi, se non addirittura risposte, proprio grazie al programma di ricerca dell'IA.

L'IA, I CASI DI SUCCESSO E LA SCOPERTA SCIENTIFICA

Remote agent è un robot che nel maggio del 1999 pianificava, in modo autonomo, a circa cento milioni di chilometri dalla Terra, la gestione delle operazioni di Deep Space One, la navicella spaziale della Nasa. Il robot generava i piani partendo da obiettivi di «alto livello», inviati da Terra, e monitorava le operazioni del veicolo spaziale durante la loro esecuzione, rilevando, diagnosticando e recuperando gli errori appena essi si presentavano. In breve, *Remote agent*, esibiva abilità di apprendimento autonomo. L'*apprendimento*, di cui è dotato *Remote agent*, sebbene risalga ad Alan Turing, si è sviluppato in modo sistematico, cioè mediante la realizzazione di artefatti, a partire dalla metà degli anni Sessanta del secolo scorso, quando cominciavano i primi successi dei *sistemi esperti*, i quali cercano di riprodurre le prestazioni di un essere umano in un determinato campo di attività. Nonostante il loro successo, ricavare le regole per la realizzazione di tali sistemi richiedeva lunghe e, a volte, inutili interviste con gli esperti di un certo dominio. Un processo come quello impiegato da *Remote agent*, invece, «apprende» le regole usando degli esempi usando l'*apprendimento automatico* o, nella sua forma più nota, *machine learning*.

Un tipico programma di *machine learning* è quello che classifica oggetti in base un insieme di attributi. Un esempio può essere la classificazione degli uccelli in «cigni» e «non cigni». Gli attributi come *taglia*, *lunghezza del collo* e *colore* possono assumere diversi valori. Ad esempio la lunghezza del collo può essere lunga, media o corta. Il colore può essere bianco o nero. E la taglia grande o piccola. Una legge o ipotesi di classificazione potrebbe essere, in questo caso, «Titty è un cigno se e solo se Titty è di taglia grande, Titty ha un collo lungo e Titty è bianco». L'ipotesi estratta può avere una forma di *generalizzazione*, come *tutti i cigni sono bianchi*, oppure di *predizione*, come il prossimo cigno sarà bianco.

Il filosofo Donald Gillies fornisce una suddivisione, nell'ambito della comunità del *machine learning*, fra due tradizioni risalenti, rispettivamente, a Alan Turing e a Herbert Simon, premio Nobel per l'economia nel 1978. Egli sostiene che la prima è caratterizzata dall'uso della logica e dall'attenzione per i problemi pratici, mentre la seconda si basa sullo studio di deduzioni. Per esempio, Turing descrive la macchina che prende il suo nome nel lavoro del 1935 quando tenta di risolvere un problema di logica astratta, e usa tale modello nel tentativo di progettare un calcolatore operativo. Simon e i suoi collaboratori, invece, hanno cominciato realizzando BACON, così denominato in onore di Francis Bacon e del suo metodo induttivo, un programma che, fra l'altro, ha riprodotto la scoperta della *terza legge del moto* dei pianeti avvenuta per opera di Keplero. Lo scienziato imperiale, usando solo due dati, la *distanza* dei pianeti dal Sole e il loro *periodo* di rivoluzione, pensò all'esistenza di regolarità (o *pattern*) legate a questi dati, trovandole dopo circa dieci anni di ricerca. BACON, il programma di Simon, con gli stessi dati di Keplero, e in modo completamente automatico, in soli pochi tentativi, ha riscoperto la terza legge del moto dei pianeti la quale, come è noto, prevede che un pianeta, nell'allontanarsi dal Sole, impiega sempre più tempo per compiere una rivoluzione completa, a prescindere dalla sua massa.

Secondo Gillies, l'approccio di Simon, sebbene abbia ottenuto ottimi risultati nella simulazione di scoperte fatte da scienziati, ha solo «riscoperto» leggi già note, mentre il *machine learning*, à la Turing, avrebbe prodotto vere *regole di generalizzazione*, o *ipotesi scientifiche*, in aree nelle quali non erano in precedenza note, come dimostra GOLEM, il sistema che nel 1992 riuscì a scoprire strutture proteiche fino allora sconosciute. Simon, nel replicare alle critiche, precisa, tuttavia, che il suo scopo è comprendere, tramite programmi come BACON, il processo cognitivo e non costruire artefatti che facciano uso di intelligenza artificiale.

L'interesse di Simon per la "riproduzione" di scoperte scientifiche, usando artefatti come programmi per calcolatore, ha dato le basi per no-

tevoli successi applicativi. Programmi come BACON sono stati di grande utilità per indagare la pratica della scoperta scientifica, oggetto della *filosofia della scienza*, impiegando metodi empirici, permettendo così di chiarire alcuni concetti fondamentali della disciplina. Questi programmi sono stati di grande utilità anche come veri e propri «scopritori», cioè scienziati, laddove la pratica scientifica avrebbe richiesto l'impiego di grandi risorse, a volte inutilmente. Un esempio è la bioinformatica, nata dall'impiego di programmi per calcolatore, tipo BACON, che sono di supporto agli scienziati per scoprire il funzionamento di processi biologici, oppure per la scoperta di nuovi farmaci.

La bioinformatica è una disciplina che, nata verso la fine degli anni Settanta del secolo scorso, applica strumenti provenienti dalla scienza dei calcolatori e dalla biologia molecolare con lo scopo di fornire modelli teorici, un altro modo per denominare le ipotesi scientifiche, in grado di spiegare fenomeni biologici. Non mancano anche obiettivi applicativi come, ad esempio, la scoperta di nuovi farmaci, come si accennava poc'anzi. Tutto ciò è facilitato dalla sempre maggiore disponibilità dei dati biologici, come quelli del genoma umano.

Un esempio è la predizione della struttura delle proteine da una sequenza di geni, un'applicazione di notevole rilievo per la funzione che alcune proteine andranno a ricoprire, soprattutto per l'eziologia di alcune gravi malattie. La funzione fisiologica della proteina, infatti, qualunque essa sia (enzima, recettore, trasportatore, proteina strutturale) dipende completamente dalla struttura tridimensionale. E alcune cause di malattia pare siano proprio da ricercarsi negli errori di *folding* (dall'inglese «ripiegamento»), cioè dalla trasformazione in struttura tridimensionale della sequenza di geni, come, ad esempio, nel caso della encefalopatia spongiforme bovina, più conosciuta con l'acronimo Bse. Ci sono anche ipotesi in merito ad altre malattie che potrebbero avere come loro causa un cattivo funzionamento di questo processo, come il morbo di Parkinson e l'Alzheimer.

Sebbene si conoscano le strutture di alcune proteine, e le relative sequenze geniche che le hanno generate, esse rappresentano solo un numero minore. Data una nuova sequenza è interessante utilizzare le sequenze per le quali già conosciamo le proteine per addestrare sistemi di apprendimento automatico in modo da scoprire nuove strutture. Golem, il sistema di apprendimento di Stephen Muggleton, introdotto sopra, nel 1992 era impegnato appunto in ricerche sulla struttura di proteine. Il problema risolto da Golem è la scoperta di strutture tridimensionali della proteina. Il sistema ha prodotto ipotesi, sotto forma di regole, per predire strutture proteiche fino allora sconosciute, mostrando, così, capacità predittive.

Oltre alle svariate applicazioni pratiche, l'apprendimento automatico,

può far luce, come sostiene anche il filosofo Donald Gillies, sul dibattito riguardante il metodo scientifico sviluppatosi tra gli «induttivisti» che fanno riferimento a Francis Bacon e i «falsificazionisti» che fanno riferimento a Karl Popper, così come sostenuto nella Sezione 6.

L'IA E IL DIBATTITO SUL «METODO SCIENTIFICO»

Come è noto, Bacone nel *Novum Organum* (1620), esponendo la sua teoria del *metodo scientifico*, nota come *induttivismo*, sostiene che uno scienziato che faccia ricorso a tale metodo colleziona una quantità accurata di osservazioni dalle quali trae delle leggi generali, o ipotesi, in grado di descrivere un certo fenomeno. Bacone chiarisce la meccanicità del suo procedimento usando la metafora del «compasso e del cerchio» secondo cui è «impossibile, anche per l'artista più dotato, tracciare un cerchio perfetto a mano libera ma con un compasso tutti possono farlo». Da questa nozione di *meccanicità* il *Lord cancelliere* propone un corollario, molto controverso, sulla natura del metodo scientifico sostenendo che grazie a esso «la scienza diventa un lavoro di routine che non ha bisogno di particolare ingegno o intelligenza».

Popper, in *Congetture e confutazioni* (1963), nega radicalmente la posizione baconiana, sostenendo che non può esistere un'osservazione pura se essa non è guidata da qualche teoria e formula una concezione alternativa del metodo scientifico, la *teoria delle congetture e delle confutazioni*, o falsificazionismo. Secondo tale punto di vista la scienza non parte da osservazioni ma da congetture, le ipotesi cui si faceva riferimento poco sopra, e compito dello scienziato è di *confutare*, o *falsificare*, le sue congetture attraverso un procedimento di critica e controllo, ad esempio facendo osservazioni o compiendo esperimenti. Una congettura che abbia resistito a un certo numero di severi controlli può essere temporaneamente accettata, con la possibilità che essa sia smentita da controlli o esperimenti successivi. Una volta confutata, la congettura va abbandonata e compito dello scienziato è di modificarla o di formularne una completamente nuova che, sottoposta a controlli ed esperimenti, deve, a sua volta, essere rifiutata se essa non li supera.

A differenza di Bacone, il quale sostiene che le teorie scientifiche possano essere ricavate da osservazioni usando un procedimento meccanico, Popper pensa che la teoria scientifica, la congettura, sia un prodotto del pensiero creativo dello scienziato e che esso non sia per nulla riconducibile a una procedura logica o automatica come sostenuto da Bacone. Le ipotesi generate da un sistema di apprendimento automatico, come Golem, e prima

di lui BACON, sembrano riconciliare le due visioni del metodo scientifico appena viste. Il sistema, infatti, genera le ipotesi meccanicamente, secondo l'auspicio baconiano. Inoltre, nel generare meccanicamente le ipotesi, il sistema ricorre continuamente a un controllo sui dati, nello spirito del falsificazionismo baconiano.

Quando Muggleton, nel 1988, aveva realizzato l'antecedente di GOLEM, nel riportare le sue prestazioni aveva già in mente il metodo adottato da GOLEM che opera, a detta del suo inventore, nello «spirito della trattazione baconiana».

EPILOGO

Le neuroscienze computazionali e l'IA, oltre a rappresentare modelli di innovazione *tout court* per l'industria del futuro, si rivelano strumenti utilissimi per affrontare e interpretare problemi sui fondamenti delle scienze che hanno stimolato l'indagine per più di venti secoli. È solo grazie a questi nuovi strumenti, infatti, che molti problemi, un tempo analizzati per vie prevalentemente concettuali, possono essere letti in chiave empirica. La presente trattazione è tutt'altro che esaustiva. Essa ha voluto mostrare, con tutti i suoi limiti, che un'indagine filosofica di tipo empirico è possibile. Sarebbe meglio dire che l'indagine filosofica rientra nell'apparato scientifico a pieno titolo.

Nella Nuova Atlantide, pubblicata postuma nel 1627, Bacone, come è noto, narra di un gruppo di naufraghi che approda sull'isola di Bensalem dove gli abitanti, cristianizzati grazie a un'arca contenente la Bibbia ma tolleranti con altre religioni, coltivano la sapienza nella *House of Solomon*, l'istituzione più importante dell'isola, realizzando esperimenti mediante il *metodo scientifico* da lui proposto nel *Novum Organum*. La popolazione bensalemita, basata su famiglia e matrimonio, secondo il racconto dei naufraghi, «conosce molte cose delle nazioni del mondo ma nessuno conosce loro». Il quadro tracciato in questa trattazione ripercorre la visione baconiana secondo cui l'indagine della natura, praticata nel *Collegio delle Opere dei Sei Giorni*, ha lo scopo di applicare la conoscenza per migliorare la società.

BIBLIOGRAFIA

- BACONE F., 1620 - *Novum Organum*. Laterza.
- DOWNNEY A.B., 2012 - *Think complexity: complexity science and computational modeling*. "O'Reilly Media, Inc."
- GILLIES D., 1996 - *Artificial Intelligence and Scientific Method*. Oxford University Press, Oxford.
- LAPLACE P.S., 1835 - *Oeuvres complètes de Laplace*. Gautier-Villars.
- LAUBE G., BERNSTEIN H.G., VEH R.W. & WEISS T., 2014 - The rationale for the localization of polyamine pathway enzymes in the brain. *The Journal of Neurobehavioral Sciences*, Üsküdar University, Üsküdar/İstanbul.
- Marr D. & Poggio T., 1976 - *From understanding computation to understanding neural circuitry*.
- MUGGLETON S., KING R.D. & STERBERG M.J.E. (1992). Protein secondary structure prediction using logic-based machine learning. *Protein Engineering*, 5 (7), pp. 647-657.
- NEWELL A., & SIMON H.A., 1972 - *Human problem solving*, 104 (9). Englewood Cliffs, NJ: Prentice-Hall.
- POPPER K.R., 1963 - *Congetture e Confutazioni*. Tr. It. Il Mulino, Bologna 1985.
- RUSSELL S.J., NORVIG P., CANNY J.F., MALIK J.M. & EDWARDS D.D., 2003 - *Artificial intelligence: a modern approach*, 2. Upper Saddle River: Prentice hall.
- SHERRINGTON C., 1910 - *The integrative action of the nervous system*. CUP Archive.
- SIMON H.A., 1965 - *Administrative behavior*, 4. Free Press, New York.
- TURING A. M., 1950 - Computing machinery and intelligence. *Mind*, 59(236), pp. 433-460.

Ernesto D'Avanzo vive a Trento ed è ricercatore all'Università di Salerno, dove insegna "Filosofia della scienza e dei processi decisionali" e "Data, Web and social analytics" presso il corso di laurea in Scienze della Comunicazione.