



I DIRITTI NELLA "RETE" DELLA RETE

Collana diretta da Franco Pizzetti

GIUSEPPE D'ACQUISTO

INTELLIGENZA ARTIFICIALE

Elementi



G. Giappichelli Editore

Capitolo Primo

Intelligenza artificiale: un primo inquadramento concettuale

SOMMARIO: 1. ... something very different ... – 2. Cose complicate e cose complesse. – 3. Ragionamento induttivo e ragionamento deduttivo. – 4. Un mondo senza volontà e con una sola rappresentazione. – 5. L'intelligenza artificiale come forma di pensiero computazionale. – 6. Quale impiego per l'intelligenza artificiale. – 7. Esercizi e verifiche di apprendimento.

1. ... something very different ...

L'intelligenza artificiale è una disciplina che ha un padre fondatore e una data di nascita certa. Il giovane padre fondatore è il matematico Alan Turing, e il momento certo della nascita è la pubblicazione nel 1950 sulla rivista *Mind* dell'articolo *Computing machinery and intelligence*. Questo il celebre l'incipit di quell'articolo in cui troviamo già molte delle questioni che successivamente la ricerca avrebbe affrontato, offrendoci il patrimonio di strumenti di cui oggi disponiamo.

I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous, if the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

.....
May not machines carry out something which ought to be described as thinking but which is very different from what a man does?

È bene soffermarsi su alcuni passaggi cruciali di queste poche righe, perché essi hanno una valenza metodologica per gli argomenti di cui ci occuperemo.

Innanzitutto, l'interrogativo “*Can machines think?*”, con l'avvento dei calcolatori elettronici che a quell'epoca iniziavano a essere studiati e costruiti, apre un ambito completamente nuovo per l'uso di una macchina, da strumento di ausilio meccanico per l'uomo a strumento di ausilio al pensiero. La macchina non è più soltanto lo strumento che l'uomo usa per costruire oggetti concreti e per trasformare lo spazio in cui vive, ma può entrare nella sfera dell'intangibile e diventare strumento per la costruzione di oggetti astratti, come un ragionamento, e per la risoluzione di problemi. Il passaggio richiede un salto concettuale molto impegnativo. Mentre infatti è evidente il tipo di ausilio meccanico che una macchina può offrirci, che di volta in volta può andare dalla capacità di sollevare oggetti sempre più pesanti, o di costruirne sempre più piccoli e precisi, o di muoversi più velocemente su distanze sempre più lunghe, l'ausilio che la macchina può offrire alla nostra capacità di pensare era, e per molti versi è tutt'oggi, assai meno evidente. Dovremmo innanzitutto essere in grado di definire cosa sia il pensiero e quali caratteristiche dovrebbe avere una macchina in grado di pensare. In questo esercizio emergono però subito due problemi. Da una parte, non disponevamo nel 1950, e in larga parte non disponiamo ancora, di una definizione oggettiva di cosa sia il pensiero dell'uomo. Dall'altra, bisogna considerare le limitazioni del linguaggio: il linguaggio naturale con il quale ci esprimiamo non sempre è idoneo a rappresentare compiutamente le funzioni svolte da un organismo complesso come un essere umano. Vi sarà sempre una zona grigia ineliminabile di ambiguità che non consentirà di catturare appieno le sfumature del nostro pensiero, che invece contano moltissimo. Tuttavia serve uno sforzo, perché rispondere all'interrogativo di Turing è quantomai necessario: l'uomo ormai da quasi un secolo intuisce questo nuovo grande ambito di applicazione per le macchine, e oggi in particolare ne avverte l'urgenza vista l'abbondanza di informazioni in cui vive e la facilità di generarle e di scambiarle su scala globale. L'informazione, potremmo dire con tutta l'ambiguità del linguaggio naturale, alimenta i nostri pensieri, e poterne generare in forma automatizzata costituirebbe per il genere umano un salto in avanti di conoscenza del mondo mai prima sperimentato. Un nuovo rinascimento, o *renAIssance*, come lo si indica, proprio con la A e la I maiuscole di Artificial Intelligence.

Alan Turing offre immediatamente una visione metodologica per uscire da questa *impasse* concettuale e linguistica. La sua idea è di abbandonare l'ambiguità del linguaggio e di concentrarci unicamente sui risultati che una macchina può produrre. Per farlo, egli assume l'esistenza di una macchina ipotetica (una macchina che in quel momento neppure esisteva) che faccia “qualcosa” per rispondere a delle domande che le sono sottoposte, in modo che l'osservatore umano le formula possa interpretare le risposte ottenute, giuste o sbagliate che siano, come il frutto di un pensiero. Una specie di imbroglio, ma che fa rientrare la questione nell'alveo tradizionale del metodo scientifico. Se questo imbroglio funziona, e l'osservatore umano non è in grado di distinguere se la risposta sia stata data da una macchina o da un uomo, tutto va come se

la macchina avesse pensato. Possiamo dire che la macchina ha prodotto un pensiero.

L'idea può essere trasformata in un esperimento ripetibile, come il metodo scientifico richiede. Il primo di questi esperimenti è il cosiddetto "imitation game", suggerito da Turing stesso, che necessita di una certa orchestrazione. Il gioco richiede la presenza di tre attori: un uomo (A), una donna (B) e un interrogatore (C). L'interrogatore non può vedere gli altri due giocatori e il suo obiettivo è identificare chi è l'uomo e chi la donna facendo delle domande ai due giocatori. Lo scopo di A è trarre in inganno C. Lo scopo di B è aiutare C nell'identificazione. Tralasciando la presenza di elementi identificativi indiretti (come il timbro della voce, che può essere dissimulato ipotizzando che il gioco si svolga con domande e risposte scritte) e gli aspetti strategici sulla condotta del giocatore A (se sia per lui più conveniente o meno dare risposte verosimili, oppure casuali), il gioco prevede che a un certo punto A possa essere sostituito da una macchina. Abbiamo così due varianti dello stesso gioco: una in cui il giocatore A è un essere umano e l'altra in cui il giocatore A è una macchina. Se la percentuale di volte in cui C dice che A è un uomo e B una donna è la stessa nelle due varianti del gioco, allora diremo che la macchina ha mostrato la stessa capacità di pensare del giocatore umano.

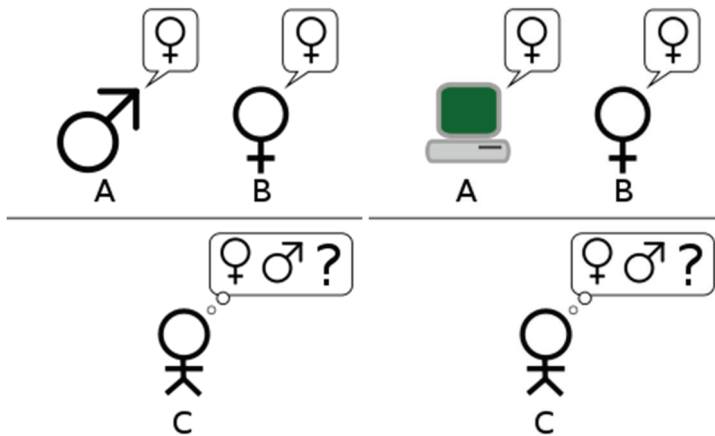


Figura 1.1: Un esempio di test di Turing

In questo modo, in assenza di un linguaggio inequivoco e di una conoscenza perfetta del funzionamento della mente umana, Turing sposta la questione semantico-linguistica (cosa è il pensiero?) associata alla domanda "Can machines think?" riportandola alla misurazione di un risultato oggettivo (quante volte l'interrogatore effettua una certa identificazione?) e alla riproducibilità dell'esperimento, ossia all'applicazione del metodo scientifico. Praticamente, il gioco prevede lo svolgimento di un certo numero N di partite in ciascuna del-

la quali l'interrogatore C pone alcune domande ad A e B. All'esito di queste partite C deve dire chi a suo giudizio è un uomo e chi una donna. Alla fine delle N partite, se l'interrogatore indovina X volte il genere del giocatore A, la percentuale di identificazioni "A è un uomo" sarà pari a X/N . Lo stesso meccanismo si applica al caso in cui il giocatore A è sostituito da una macchina. In questo caso, l'interrogatore effettuerà l'identificazione "A è un uomo" Y volte su N tentativi. Si dirà che la macchina avrà superato il test di Turing, e dunque che la macchina avrà "pensato", se le due percentuali sono molto simili, ovvero se $X/N \approx Y/N$. Il gioco, come ovvio, è totalmente simmetrico e i ruoli di A e B sono interscambiabili. Dunque, per l'esito del gioco, il genere dei giocatori A e B non conta, purché questo sia diverso (o purché i due giocatori abbiano una diversa caratteristica in grado di distinguerli chiaramente). Turing non dice come deve essere fatta una macchina pensante, ma soltanto i requisiti che l'esperimento dovrà avere per poter affermare che una macchina sia stata in grado di elaborare un pensiero.

A partire dall'intuizione di Turing, molte varianti del gioco sono state ipotizzate. La più rilevante è quella avanzata da Joseph Weizenbaum negli anni '60 del novecento realizzata attraverso il programma Eliza, una versione del gioco di Turing con un solo interrogatore umano e una macchina. Il nome del programma dice molte cose sul tipo di gioco: esso trae spunto dalla commedia *Pygmalion* di George Bernard Shaw, nella quale una donna del popolo, la fioraia Eliza Doolittle, attraverso i precetti di un istruttore, il prof. Henry Higgins, viene trasformata in una raffinata dama della *high society* dell'epoca. Infatti, il programma Eliza è stato realizzato per conversare con un interrogatore umano, in modo da dissimulare la sua natura automatizzata, e da imparare nel tempo a rispondere alle domande che gli venivano poste in modo sempre più indistinguibile da un giocatore umano; proprio come Eliza, la protagonista della commedia, aveva appreso i modi aristocratici del suo tempo. In questo caso Weizenbaum, a differenza di quanto fatto da Turing, offre anche delle indicazioni progettuali sulle caratteristiche che dovrà avere una macchina pensante. Il seguente frammento di un ipotetico dialogo tra Eliza, che interpreta il ruolo di uno psicoterapeuta, e una paziente (l'interrogatore) può fare capire l'artificio usato all'epoca per riuscire a ingannare l'interrogatore.

Interrogatore: Gli uomini sono tutti uguali.

Eliza: In che senso?

Interrogatore: In un modo o nell'altro ci tormentano sempre.

Eliza: Può farmi un esempio?

Interrogatore: Be', è stato il mio ragazzo a farmi venire qui.

Eliza: È stato il suo ragazzo a farla venire qui.

Interrogatore: Dice che sono sempre depressa.

Eliza: Mi dispiace sentire che è depressa.

Interrogatore: È vero. Sono infelice.

Eliza: *Pensa che venire qui l'aiuterà a non essere infelice?*
Interrogatore: *Ho bisogno di aiuto, questo è sicuro.*
Eliza: *Che cosa cambierebbe per lei se ricevesse aiuto?*
Interrogatore: *Forse imparerei a sopportare mia madre.*
Eliza: *Mi dica qualcosa della sua famiglia.*
Interrogatore: *Mia madre si occupa di me.*
Eliza: *Chi altri della sua famiglia si occupa di lei?*
Interrogatore: *Mio padre.*
Eliza: *Suo padre.*
Interrogatore: *Per certi versi lei somiglia a mio padre.*
Eliza: *Che somiglianze vede?*
Interrogatore: *Lei non è molto aggressivo, ma penso che non voglia farmelo capire.*
Eliza: *Perché pensa che non sia molto aggressivo?*
Interrogatore: *Lei non litiga con me.*
Eliza: *Perché pensa che io non litighi con lei?*
Interrogatore: *Perché ha paura di me.*
Eliza: *Le fa piacere pensare che ho paura dei lei?*
Interrogatore: *Mio padre ha paura di tutti.*
Eliza: *Che cos'altro le viene in mente quando pensa a suo padre?*
Interrogatore: *I prepotenti.*
Eliza: *Questo ha a che fare con il fatto che il suo ragazzo l'ha fatta venire qui?*

Il “*something very different*” che ci inganna, e che ci può indurre a credere che il programma Eliza stia realmente pensando le risposte che fornisce all'interrogatore, consiste in questo caso nella ricerca da parte del programma di alcune parole chiave nelle domande formulate dal paziente. Se queste vengono trovate, il programma trasforma la frase secondo una regola prefissata. Ad esempio, ribaltando un'affermazione in una domanda di approfondimento, come nel passaggio

Interrogatore: *Forse imparerei a sopportare mia madre.*
Eliza: *Mi dica qualcosa della sua famiglia.*

in cui la parola chiave è “*madre*” a cui si associa il termine “*famiglia*”, e la radice della domanda di approfondimento è “*mi dica qualcosa della sua*”. Altrimenti, se non è possibile individuare parole chiave note, il programma ripete in altro modo l'affermazione precedente dell'interrogatore, come in

Interrogatore: *Dice che sono sempre depressa.*
Eliza: *Mi dispiace sentire che è depressa.*

in cui la parola chiave è “*depressa*” e la frase che viene usata in testa alla replica è “*Mi dispiace sentire che è*”. Oppure ancora, sempre in assenza di parole chiave, il programma genera un'osservazione priva di contesto per fare ripartire il dialogo, come in

Interrogatore: *In un modo o nell'altro ci tormentano sempre.*

Eliza: *Può farmi un esempio?*

Lo schema logico su cui si articola il funzionamento del programma Eliza è rappresentato in figura 1.2, in cui si notano le varie tabelle che il programma usa per individuare le parole chiave pronunciate dal paziente, e per associarle a possibili frasi di risposta. Si noti il ruolo di “pigmalione” svolto dagli utenti stessi (i pazienti di Eliza), che nel tempo con le loro domande possono alimentare la memoria di Eliza con nuove espressioni, che potranno essere impiegate per consentire al programma di rispondere in modo sempre più assimilabile ad un vero terapeuta umano. La macchina non soltanto “pensa”, ma anche “impara”. Non c’è naturalmente niente nel funzionamento di Eliza che possa essere considerata una forma di intelligenza. È un inganno. Ma questo stesso meccanismo di inganno è quello su cui si basano oggi i programmi dialogici automatizzati, o chatbot.

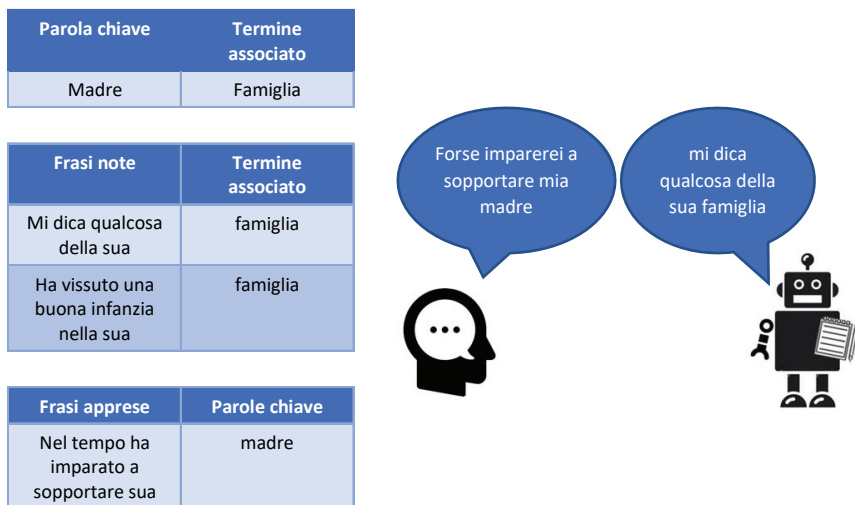


Figura 1.2: Schematizzazione di un programma dialogico

2. Cose complicate e cose complesse

Da allora molte diverse forme di pensiero sono state ipotizzate per una macchina, non soltanto basate sul paradigma della dissimulazione di Turing. Anzi, lo schema del test di Turing, che rimane tuttora un validissimo indicatore del livello di “abilità” della macchina nell’ingannare l’uomo, è stato progressivamente abbandonato a favore di un approccio più moderno all’intelligenza artificiale, che potremmo definire di tipo *problem solving*. Per dirla

con Stuart Russel e Peter Norvig: “*planes are tested by how well they fly, not by comparing them to birds*”. Ossia, una volta compreso che le machine sono in grado di produrre un certo tipo di risultato che può essere assimilato a una forma di pensiero, non serve concentrarsi sull’efficacia dell’inganno, effettuando questa comparazione continua tra ragionamento umano e ragionamento automatico, ma è meglio rivolgere l’attenzione a quella modalità di pensiero così differente dalla nostra che le macchine sono in grado di realizzare e svilupparla per la risoluzione di problemi specifici. Negli ultimi decenni, dunque, la ricerca scientifica ha di fatto smesso di considerare l’intelligenza artificiale come un tentativo di riprodurre il modo in cui l’intelligenza umana ragiona, e si è maggiormente concentrata sulle svariate tecniche anche molto diverse tra loro (e questa assenza di unitarietà porta molti osservatori a sostenere che l’intelligenza artificiale sia più un’arte più che una scienza), che applicate alla risoluzione di determinati problemi producono dei risultati assimilabili al modo in cui un essere umano avrebbe risolto quel problema. In altri termini, se un’intelligenza umana disponesse delle risorse di calcolo della macchina, e se si applicasse alla risoluzione di quei problemi, la modalità con cui perverebbe alla loro soluzione sarebbe la stessa di quella individuata dalla macchina. Il punto è che l’intelligenza dell’uomo molto difficilmente si applicherebbe alla risoluzione di quei problemi, perché se osservati dal punto di vista di una intelligenza creativa, come quella umana, essi appaiono lunghi, complicati e talvolta persino noiosi.

Per iniziare a comprendere il tipo di intelligenza non creativa che la macchina può effettivamente esprimere, il suo ambito di applicazione e la differenza principale con l’intelligenza umana, può essere utile fare riferimento alla risoluzione di un rompicapo come il cubo di Rubik. Su Youtube ci sono molti video di ragazzi che sono in grado di risolvere il rompicapo in modo prodigioso e in situazioni anche molto disagiati (sott’acqua, ad esempio, o lanciandosi da un aereo con il paracadute, o addirittura risolvendone tre contemporaneamente facendoli volteggiare in area come farebbe un giocoliere!). Si tratta certamente di ragazzi che nel risolvere il rompicapo esprimono una particolare forma di intelligenza. Ma probabilmente non passeranno alla storia per questo loro talento. Difficilmente per quel tipo di attitudine spenderemmo l’aggettivo intelligente. Nell’ambiguità del nostro linguaggio naturale, di solito l’intelligenza ha più a che fare con la capacità umana di far volgere il corso degli eventi a proprio favore, o con la presenza di un guizzo creativo nella realizzazione di un’opera, anche se manifestato una volta soltanto ma decisiva.

L’attitudine creativa dell’intelligenza umana nel far volgere il corso degli eventi a proprio favore, o nella realizzazione di un’opera, non è schematizzabile in sequenze di istruzioni logico-matematiche; non è un algoritmo. La risoluzione di un rompicapo invece lo è. E i passi di queste sequenze logico-matematiche necessari per risolvere problemi rompicapo possono essere molti e molto articolati, in modo da renderli praticamente intrattabili da parte dell’intelligenza umana, per varie ragioni connesse alle nostre limitate risorse compu-

tazionali e di memoria. Eppure sulla risoluzione di problemi rompicapo che richiedono la metodica applicazione di sequenze di istruzioni logico-matematiche si basa oggi il successo di molte iniziative industriali in diversi settori. Ecco che la prospettiva astratta di Turing “*Can machines think?*” diventa molto concreta. La macchina non ha, e sempre meno avrà, problemi a trattare grandi moli di dati e ad eseguire sequenze molto lunghe di istruzioni. E lo farà in tempi sempre più rapidi. In questo modo potrà estendere la nostra capacità umana di risoluzione dei problemi ad un’ampia categoria di situazioni per le quali in assenza di macchine intelligenti ci arrenderemmo, perché il numero di combinazioni di casi da valutare sarebbe per noi intrattabile.

Un esempio per tutti: il gioco degli scacchi. Esso appartiene alla categoria dei cosiddetti “giochi parzialmente risolti”, nel senso che il numero di partite che possono essere giocate su una scacchiera convenzionale 8x8 non è noto esaustivamente e dunque non è possibile individuare a priori una strategia infallibile. Claude Shannon, un altro dei pionieri dell’intelligenza artificiale (e padre fondatore di un’altra scienza: la teoria dell’informazione), in un lavoro del 1950, *Programming a Computer for Playing Chess*, apparso sulla rivista *Philosophical Magazine*, stimò per difetto che esistessero circa 10^{120} possibili partite a scacchi costituite da 40 coppie di mosse da parte dei due giocatori. Si tratta di un numero incredibilmente elevato. Per averne un’idea, si consideri che il numero stimato di atomi presenti nell’universo è nell’ordine di 10^{80} , ossia circa 40 ordini di grandezza in meno di 10^{120} . Questo significa che, se ogni atomo dell’universo fosse un minicomputer in grado di eseguire mille miliardi di istruzioni al secondo (10^{12} istruzioni ogni secondo) e tutti questi computer fossero interconnessi tra loro, ci vorrebbero circa 3 miliardi di miliardi di secoli prima che il gioco degli scacchi diventi un gioco totalmente risolto! Non avete letto male. Gli scacchisti possono stare tranquilli: c’è ancora un po’ di tempo prima che l’intelligenza artificiale (qualsiasi futura intelligenza artificiale) renda il loro gioco totalmente prevedibile. L’area dei problemi come questo degli scacchi, ossia con un numero di combinazioni intrattabile da un essere umano, è una di quelle in cui l’intelligenza artificiale potrà trovare grandi applicazioni.

Il gioco è una costante nella moderna ricerca scientifica nel settore dell’intelligenza artificiale, perché nel gioco sono concentrate spesso tutte le proprietà del problema specifico che si vuole risolvere, e perché la possibilità di rappresentare il problema come un gioco consente anche di ricondurlo a categorie riconoscibili e concettualmente trattabili dalla nostra intelligenza. Le regole per la soluzione di quel gioco poi, nel caso concreto, potranno consentire di realizzare applicazioni di successo dell’intelligenza artificiale nel campo della robotica, o della guida assistita o della logistica, giusto per citare alcuni esempi. Non si sottovaluti dunque l’importanza del gioco, a cui faremo riferimento molte volte in questo libro.

Alla luce di questi semplici esempi, guardando soltanto ai risultati che si possono conseguire come frutto di una qualsiasi forma di intelligenza, e senza

richiedere la definizione di cosa essa sia, si delinea una prima classificazione delle applicazioni per le quali un'intelligenza, umana o artificiale, può essere impiegata. Ci sono azioni complesse, in quanto creative, che sono tipiche dell'intelligenza umana. Ci sono azioni complicate, perché ad esempio richiedono di considerare molte varianti, che possono anche essere effettuate da una macchina, e che anzi da un certo punto in poi possono essere effettuate soltanto da una macchina. E poi ci sono le azioni più semplici. Tra queste ultime, alcune sono più creative, e sono trattabili da un essere umano, altre invece sono di tipo esecutivo e possono essere effettuate anche da una macchina. Possiamo immaginare di rappresentare queste applicazioni all'interno di quattro quadranti come quelli mostrati in figura 1.3.

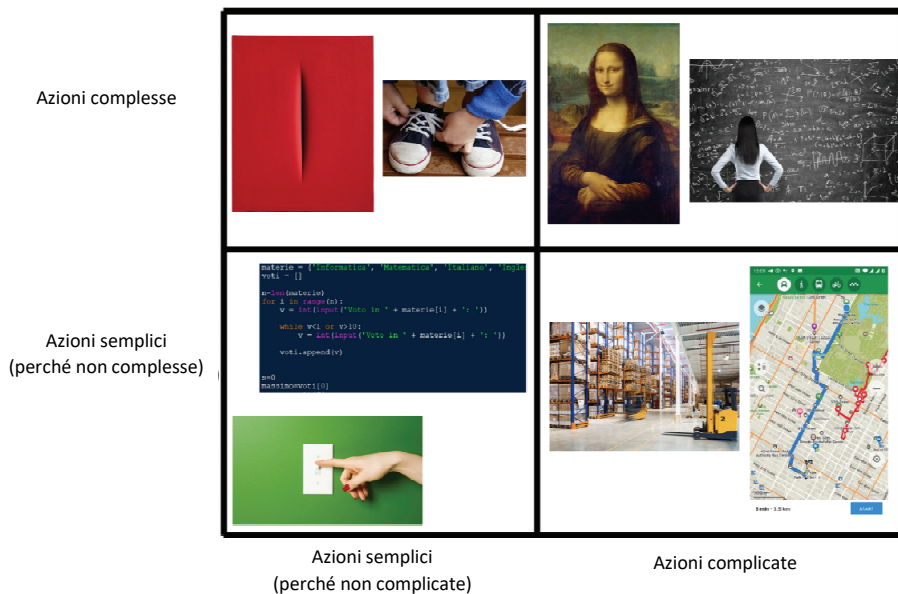


Figura 1.3: Azioni complicate e azioni complesse

Il quadrante in basso a sinistra è popolato da applicazioni molto semplici, in quanto non complesse né complicate. Sono azioni che possono essere impartite come comandi, che non richiedono particolari riflessioni e che possono indifferentemente essere eseguite da un essere umano o da una macchina, come azionare un interruttore. Anche l'esecuzione di un programma informatico rientra in questa categoria. In alto a sinistra ci sono azioni non complicate, ma che richiedono comunque una riflessione e una complessa coordinazione tra pensiero e azione, tra mente e corpo, come allacciare le stringhe delle scarpe. A questo quadrante appartengono anche manifestazioni dell'ingegno apparentemente semplici (non complicate) ma molto concettuali (ossia complesse),

come gli squarci su una tela di Lucio Fontana o le opere dada di Marcel Duchamp. Sono tutte attività tipicamente umane. Nella parte di destra abbiamo in alto tutte le opere dell'ingegno che richiedono una tecnica per la quale è richiesto un lungo addestramento, come la letteratura, l'arte, la scienza. Anche questi sono frutti dell'intelligenza dell'uomo. Il basso a destra abbiamo invece un'area in grande espansione, della quale oggi vediamo le prime applicazioni concrete, in cui la macchina tratta enormi quantità di dati per risolvere concreti problemi, come l'individuazione del migliore percorso all'interno di una città, oppure la ripartizione di carichi di lavoro in una catena di produzione o la distribuzione di merci. Questo è l'ambito in cui l'intelligenza artificiale potrà trovare molteplici applicazioni in futuro. Attualmente non sappiamo prevedere fino in fondo i benefici che verranno per il genere umano da queste applicazioni, intuiamo soltanto che vi sono spazi rilevanti e promettenti di intervento in molti settori dalla medicina alla logistica, dai trasporti all'agricoltura.

Oggi l'intelligenza artificiale offre moltissime tecniche per trattare dati in grande quantità e per risolvere molti problemi logici e combinatori specifici. Eppure si tratta di una disciplina che non ha un carattere di unitarietà. Essa poggia in effetti su scienze più consolidate e tradizionali, come la matematica, la logica e il calcolo delle probabilità per produrre i propri risultati. Potremmo dire che l'intelligenza artificiale è una combinazione di queste scienze, applicata alla risoluzione di specifici problemi. In questo senso è una disciplina che sarà sempre in divenire, dal momento che non possiamo prevedere il tipo di problemi che le esigenze dell'uomo genereranno. Da un punto di vista strettamente logico-matematico, il risultato a cui si giunge con tecniche di intelligenza artificiale è chiaro e, di volta in volta, come vedremo, potrà consistere in una deduzione logica, un'ottimizzazione, una previsione, una classificazione. Quale sia il risultato in termini di conoscenza del mondo che traiamo dall'analisi di queste moli di dati attraverso tecniche di intelligenza artificiale non è però altrettanto chiaro. Vi è certamente la scoperta di correlazioni non immediatamente visibili, e una migliore sistematizzazione delle nostre categorie di pensiero, per via dell'accresciuta capacità data dalle macchine di individuare similarità e differenze. Alcuni sostengono che sia anche possibile progettare intelligenze artificiali in grado di imparare e che per effetto di questo apprendimento le macchine potranno scoprire nuove regole a cui sono soggetti i fenomeni fisici e socio-economici, mai prima immaginate. Ma forse anche per le macchine c'è un orizzonte non raggiungibile, e molte autorevoli voci della comunità scientifica, attualmente impegnate nello sviluppo di tecniche di intelligenza artificiale e capaci di fornire riflessioni filosofiche sul tema, sono per esempio estremamente scettiche sulla possibilità che le macchine possano mai individuare la causa ultima dei fenomeni, o condurre analisi controfattuali (del tipo "cosa sarebbe successo se"), che richiederebbero quel grado di creatività tipico degli uomini nel momento in cui effettuano una scoperta, non schematizzabile da nessun tipo di regola logico-matematica oggi esistente né prevedibile.

Naturalmente, l'ingresso della macchina nell'ambito del pensiero, anche soltanto per lo svolgimento di azioni complicate (ma non complesse), pone molti interrogativi sulla coesistenza tra uomini e macchine. Il tema è oggi ampiamente dibattuto e ogni ipotesi è aperta. Di certo, la macchina offre all'uomo la possibilità di risolvere problemi prima intrattabili, per via dell'esplosione combinatoria delle varianti a cui essi danno luogo, allargando progressivamente l'insieme dei problemi risolvibili. Non è però affatto detto che nel fare questo la macchina invada l'ambito del pensiero umano, come i pessimisti ritengono. Il problema non è tanto se la macchina possa pensare come l'uomo, sottraendo all'uomo spazi di autodeterminazione nella produzione creativa, congettura che andrà verificata, ma se l'uomo sarà costretto per effetto della presenza di macchine (a loro modo) intelligenti a ragionare come la macchina e in questo a vedere limitata la propria libertà. Guardando alla storia dell'uomo, che in questi aspetti presenta tratti ricorrenti, non è irragionevole ipotizzare che l'uomo, nel disporre di macchine in grado di ampliare l'insieme dei problemi che possono essere risolti con tecniche logico-matematiche, finisca per creare nuove forme di divisione tra uomini e uomini. Tra gli scenari futuri da considerare, determinati dalla presenza di macchine intelligenti, vi è senz'altro la creazione di nuove classi sociali, forse anche più nettamente separate di quanto non siano già attualmente: da una parte alcuni uomini che potrebbero trovare il loro "posto sociale" nello spazio creato dal funzionamento delle macchine intelligenti, diventando in un certo senso strumento della macchina (ad esempio, alimentando le macchine con attività massive di data entry, come avviene già oggi per le applicazioni social, oppure curando il buon funzionamento degli algoritmi), dall'altra altri uomini che avranno le capacità e la forza di sottrarsi a questa attrazione e che potranno utilizzare la macchina come un loro strumento, delegandole la fatica delle azioni più ripetitive e noiose.

Proprio questo è il compito del decisore, ad ogni livello in cui opererà: individuare regole che siano capaci di trarre tutto il meglio dall'intelligenza artificiale, e ve ne sono tutte le premesse, e di scongiurarne gli effetti negativi, offrendo a tutti un accesso universale e non discriminatorio ai servizi e ai benefici attesi, promuovendo la competizione e la fiducia delle più giovani generazioni nella formazione e nel progresso tecnologico. Prendere il meglio dall'intelligenza artificiale e allontanare il peggio. È un compito estremamente complesso che spetta in particolare agli operatori del diritto, ma non soltanto a questi. Esso investe infatti tutte le anime delle nostre società e richiederà le migliori idee di scienziati, filosofi, economisti. Per giungere a regole e tutele efficaci bisogna però conoscere intimamente i presupposti tecnici su cui funziona l'intelligenza artificiale. Il regolatore in un mondo popolato da macchine intelligenti non potrà essere soltanto uno spettatore.

3. Ragionamento induttivo e ragionamento deduttivo

È opportuno a questo punto accennare alle due forme di ragionamento che ci consentono di accrescere la nostra conoscenza sui fenomeni del mondo o, meglio, di ridurre la distanza tra il grado di incertezza iniziale con cui affrontiamo un problema e il grado di incertezza finale con cui giungiamo a una conclusione: il ragionamento induttivo e il ragionamento deduttivo. Esse valgono anche per il frutto del pensiero di una macchina, qualsiasi sia la forma in cui la macchina saprà esprimerlo.

Il ragionamento induttivo è un procedimento che cerca di stabilire una legge universale partendo da casi particolari. Esso si basa sull'esperienza ed è una forma di inferenza ampliativa, nel senso che ci consente di fare un passo in più rispetto a quanto osservato. Un esempio canonico del procedimento induttivo è il seguente:

Ho visto un primo corvo ed era nero
Ho visto un secondo corvo ed era nero
Ho visto un terzo corvo ed era nero

Ho visto il centesimo corvo ed era nero

Finora ho visto soltanto corvi neri
Dunque
Tutti i corvi sono neri

È di tutta evidenza, sulla scorta di questo semplice esempio, come ogni legge induttivamente giustificata sulla base dell'esperienza abbia in realtà una ineliminabile radice soggettiva. Cosa ci autorizza a ritenere che, anche dopo la centesima o millesima osservazione, siamo in presenza di una legge generale? L'idea di ritenere collegate tra loro le premesse (“*ho visto un primo corvo ed era nero*”; “*ho visto un secondo corvo ed era nero*”; e così via) e la conclusione del ragionamento (la parte del discorso dopo il “*dunque*”, cioè la legge generale “*tutti i corvi sono neri*”) nascerebbe soltanto da un istinto, dall'abitudine di vedere le premesse accadere in una certa sequenza. Con il tempo, e in particolare con le riflessioni di David Hume a partire dal '700, la pretesa capacità ampliativa del ragionamento induttivo è stata mitigata con l'introduzione della nozione di probabilità, su cui torneremo, a favore di una lettura più umile dei fenomeni del mondo. Sempre con riferimento all'esperimento del corvo, ciò che in effetti ci è lecito affermare è che più l'esperienza procede, minore è il livello di incertezza rispetto alla prossima osservazione. In questo senso, un procedimento induttivo più razionale potrebbe essere il seguente:

Ho visto un primo corvo ed era nero, probabilmente il secondo sarà nero

Ho visto un secondo corvo ed era nero, ancor più probabilmente il terzo sarà nero

Ho visto un terzo corvo ed era nero, ancor più probabilmente il quarto sarà nero

.....

E così via, in un procedimento di ampliamento parziale ma progressivo della conoscenza che non ha e non può avere fine. Il livello di fiducia sulla prossima osservazione cresce ad ogni osservazione. Dopo N osservazioni di corvi neri, ciò che potrà succedere alla prossima osservazione è che per la prima volta si manifesti un corvo non nero, ma se ciò accadrà io avrò avuto N casi favorevoli (tutte le volte in cui ho osservato un corvo nero), su un numero $N+1$ di casi osservati (tutti quelli in cui ho osservato un corvo nero, a cui si aggiunge il nuovo caso, incognito, in cui potrei anche osservare un corvo non nero). Il mio livello di fiducia induttivo sulla conoscenza di quel fenomeno alla N -esima osservazione, e prima che si compia la $(N+1)$ -esima, sarà dunque dato dal rapporto $N/N+1$ che è un numero più piccolo di 1 che cresce ad ogni osservazione, segno che non sarò mai in grado di raggiungere la conoscenza perfetta ma, nella migliore delle ipotesi, soltanto di ridurre progressivamente il mio stato di incertezza.

Occorre poi rilevare che l'approccio induttivo non è mai puramente empirico, ossia basato integralmente su osservazioni, e che c'è sempre una componente aprioristica che rende in effetti l'osservazione intrinsecamente non neutrale. In ogni approccio asseritamente empirico e induttivo, l'uomo tende sempre, inconsciamente o per interesse, a sovrapporre i propri schemi mentali alla realtà che osserva. C'è sempre una teoria, o una metafisica, dietro la ricerca di una teoria. Già soltanto scegliere cosa osservare è una scelta non induttiva carica di teoria. Per rendersene conto, rimanendo in tema di corvi, immaginiamo questo piccolo esperimento.

Ipotizziamo l'esistenza di un mondo composto soltanto da due tipi di "oggetti": alcuni corvi, dei quali non sia specificato il colore, e alcuni "oggetti" rossi, tra i quali potrebbe esservi una rara specie di corvi rossi. In particolare, immaginiamo di avere una conoscenza parziale di questo mondo e di volerlo conoscere meglio per via induttiva. Ad esempio, immaginiamo di sapere che in questo mondo sono presenti N corvi, dei quali non so il colore, e M oggetti rossi, tra i quali potrebbero esservi dei corvi.

La domanda è: in questo piccolo mondo, come conviene procedere per essere certi dell'affermazione "tutti i corvi sono neri"? Meglio andare in cerca dei corvi, o puntare agli oggetti rossi? Questa è una situazione più semplice del caso generale, in quanto siamo in presenza di un mondo chiuso con un numero finito di possibilità, quindi si può raggiungere la certezza su una legge generale. A voler considerare tutte le opzioni, infatti, da una parte c'è la separazione completa tra l'insieme dei corvi neri e l'insieme degli oggetti rossi, dall'altra abbiamo invece due possibili casi di compenetrazione tra i due insiemi:

il caso in cui $N > M$ e dunque avremo N corvi, M dei quali sono rossi e gli altri $N - M$ neri, oppure il caso in cui $N < M$ e avremo M oggetti rossi, N dei quali sono dei corvi. In mezzo, tutte le altre possibilità. Alcune di queste opzioni sono rappresentate in figura 1.4. In effetti, le due modalità empiriche per giungere alla nostra legge generale, andare in cerca dei corvi sperando che tutti siano neri, oppure cercare gli oggetti rossi sperando di non trovarvi dei corvi, sono equivalenti in vista dell'obiettivo. Scegliere una strada o l'altra è però un meta-problema, che si sovrappone al nostro esperimento e che può condizionarlo. Molti infatti sono i fattori che possono influenzare la condotta del ricercatore: il tempo disponibile, le risorse che possono essere impiegate per la ricerca, il costo della ricerca, un interesse personale. Questi fattori, in ragione del contesto, possono intervenire in modo assai vario, non sempre schematizzabile logicamente, e dare luogo a scelte anche controintuitive (si veda al riguardo l'esercizio 5 al termine di questo capitolo). Non c'è dunque soltanto l'osservazione empirica e l'esperienza. Il metodo conta, ma oltre l'esperienza c'è sempre l'intuito, che non è formalizzabile in un algoritmo.

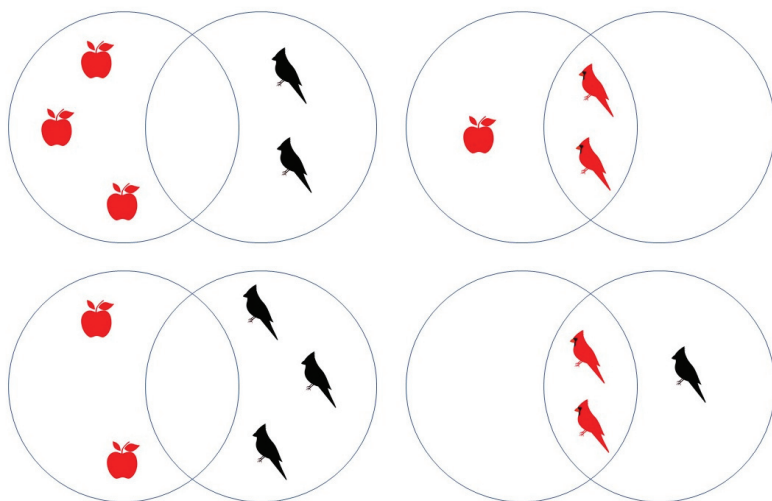


Figura 1.4: Mondi con tre oggetti rossi e due corvi (in alto), e mondi con due oggetti rossi e tre corvi (in basso)

Contrapposto al ragionamento induttivo è il ragionamento deduttivo, che è il procedimento razionale mediante il quale è possibile derivare una certa conclusione a partire da premesse più generali, dentro cui quella conclusione è implicita. Il metodo deduttivo parte sempre da un insieme di postulati, che non hanno bisogno di essere dimostrati e che costituiscono la condizione di avvio del ragionamento. A partire da questi postulati, seguendo una serie di concatenazioni logiche si dimostra che una certa affermazione è conseguenza