



Daniele Missiroli

Sono nato a Ravenna e vivo a Bologna, dove lavoro come analista programmatore e consulente privacy nell'azienda che ho fondato. Laureato in fisica con lode, sono attratto da numeri, astronomia e scienze in generale, oltre che da fantascienza, fumetti e film. Scrivo libri di fantascienza e ho creato una serie ambientata su un lontano pianeta. Ho scritto anche un libro di matematica sul calcolo mentale.

Costruiamo una “Intelligenza Artificiale”

Oggi si sente molto parlare di **Intelligenza Artificiale** (indicata spesso con la sigla **IA**). Giornali, televisione e anche diversi film la citano, e a volte la mostrano, come robot (o androidi o computer) che ragionano come gli esseri umani. Proviamo a costruirne logicamente una?

La prima cosa che si deve assolutamente sapere è che la IA è un **programma**. La differenza sostanziale fra un qualsiasi dispositivo elettronico e un computer è la presenza di un oggetto che viene chiamato CPU (**Unità Centrale di Processo**). Una lampadina può accendersi se un essere umano preme un pulsante. Una lavatrice può lavare i panni se un essere umano gira una manopola e preme dei pulsanti. Se un dispositivo contiene al suo interno una CPU, è un computer. È un computer anche se non è dotato di tastiera, mouse e video.

Ebbene sì: sono le CPU a fare la differenza. Ne esiste solo una? No, da quando l'Intel ha costruito la prima, ne hanno fabbricate centinaia di modelli. Le aziende che oggi sono in grado di produrle si contano sulle dita di una mano. Credo siano rimaste solo in quattro, ma ogni azienda produce centinaia di modelli. Le ultime generazioni sono le più potenti e le più costose. Una CPU può arrivare a costare anche 500 dollari o più. Un computer con una di queste CPU, quindi, può costare migliaia di dollari. Le vecchie CPU, al contrario, possono costare anche solo pochi dollari, ed è per questo che oggi conviene montare una CPU anche in dispositivi a basso costo, invece di inserirvi dell'elettronica in grado di svolgere una sola funzione. Se costruisco un frullatore con una manopola, girando la quale scelgo diversi tipi di rotazione delle lame, quando costruirò un modello nuovo, dovrò rifare tutto. Se invece uso una CPU, nel modello nuovo metterò gli stessi componenti elettronici con un programma nuovo, che farà cose diverse.

Rispetto agli elementi elettronici normali, una CPU può eseguire istruzioni. Istruzioni semplici, banali, minimaliste, ma la sua velocità si misura in miliardi di istruzioni al secondo, per cui, in poco tempo, è in grado di svolgere compiti anche molto complessi. Prendiamo uno schermo che abbia una risoluzione di 2000 pixel per 1000. In totale ci sono due milioni di puntini. Una CPU può muoverli tutti, uno per uno, a destra di un pixel e ripetere l'operazione fino a farci percepire il movimento dell'immagine. Ecco che abbiamo realizzato un programma che ci permette di vedere un film. Quando lo guardiamo, ricordiamoci che la CPU sta muovendo i puntini uno per uno. Ne muove miliardi in un secondo e noi vediamo il film.

Ora che abbiamo identificato la componente **fisica** di una potenziale IA, chiediamoci se è sufficiente dargli corrente per farla funzionare. La risposta è no!



Daniele Missiroli

Sono nato a Ravenna e vivo a Bologna, dove lavoro come analista programmatore e consulente privacy nell'azienda che ho fondato. Laureato in fisica con lode, sono attratto da numeri, astronomia e scienze in generale, oltre che da fantascienza, fumetti e film. Scrivo libri di fantascienza e ho creato una serie ambientata su un lontano pianeta. Ho scritto anche un libro di matematica sul calcolo mentale.

Abbiamo la nostra CPU, le diamo corrente, lei si accende e cosa fa? Niente! La situazione di quella CPU è analoga a un tetraplegico cieco e sordomuto steso su un letto. Senza un programma, la CPU non può fare niente. Ecco perché, accendendo un computer, esiste un chip che contiene il programma di base, un meccanismo automatico lo carica nella memoria centrale e la CPU inizia a eseguirlo. Se quella fase fallisse, non vedremmo nemmeno la scritta “Errore” sul video. La CPU non sa che esiste un video. Non sa di quanta memoria dispone, non sa niente. È come se fosse nata nel momento dell'accensione. Ecco perché alcuni programmatori non spengono mai il proprio computer. Spegnerne un computer è come ucciderlo.

Ora abbiamo capito che per costruire una IA abbiamo bisogno di una macchina dotata di una CPU (compreremo la più potente, tanto questo è solo un esercizio logico) e dobbiamo scrivere un programma. Che cosa dovrà fare questo programma per farci credere che la macchina sia “intelligente”? Immaginiamo di dotarla di infinita memoria, di microfono e di altoparlanti, così sarà più facile interagire con lei. Le parole italiane che conosce una persona dotata di un certo livello culturale sono meno di 20.000. Usiamone 50.000 e con queste parole costruiamo tutte le domande possibili. Associamo a ogni domanda una risposta e memorizziamo tutto nell'infinita memoria della nostra macchina. Ora chiediamo a un essere umano (che chiameremo Danny) di fare delle domande alla nostra macchina. Se Danny crederà di aver parlato con una persona, il test sarà stato superato e la macchina sarà dichiarata essere una vera Intelligenza Artificiale. Naturalmente Danny cercherà di porre domande “trappola” per far cadere in contraddizione la nostra macchina.

Danny decide di iniziare con dieci domande identiche. Alla fine ci dice: “Ho parlato con una macchina, perché risponde sempre nello stesso modo!”. Subito il programmatore modifica il programma. Ora la macchina ha tante risposte corrette per ogni domanda, non una sola. Rifacciamo il test.

Danny decide di porre cento domande identiche. Alla fine ci dice: “Ho parlato con una macchina! Risponde in modo differente alla stessa domanda, ma risponde sempre! Un essere umano si stuferebbe e mi manderebbe al diavolo!”. Il programmatore modifica di nuovo il software: ora la macchina conta le domande uguali e, se le sono troppe, dopo la terza manda al diavolo l'interlocutore. Rifacciamo il test.

Danny decide di porre dieci domande diverse, ma ognuna la pone dieci volte. Alla fine ci dice: “Ho parlato con una macchina! Dopo la terza domanda identica mi manda al diavolo, ma lo fa sempre dopo la terza. Un essere umano non è così regolare”. Il programmatore modifica di nuovo il programma: ora la macchina conta le domande uguali, genera un numero casuale e, se le vengono poste troppe domande uguali, dopo quel numero casuale, manda al diavolo Danny. Non sarà mai più dopo la terza in modo fisso. Rifacciamo il test.

Danny decide di porre cento domande diverse, ma ognuna la pone dieci volte. Alla fine ci dice: “Ho parlato con una macchina! Ora si stanca della stessa domanda in maniera casuale, ma



Daniele Missiroli

Sono nato a Ravenna e vivo a Bologna, dove lavoro come analista programmatore e consulente privacy nell'azienda che ho fondato. Laureato in fisica con lode, sono attratto da numeri, astronomia e scienze in generale, oltre che da fantascienza, fumetti e film. Scrivo libri di fantascienza e ho creato una serie ambientata su un lontano pianeta. Ho scritto anche un libro di matematica sul calcolo mentale.

ho fatto la domanda 100 volte e qualcosa mi ha sempre detto; a volte una risposta corretta, a volte un insulto, ma un essere umano smette di parlare con me dopo un po', perché si accorge che la conversazione non ha senso". Il programmatore aggiunge un'altra funzione: ora la macchina ha anche la capacità di non rispondere e addirittura di riprendere la conversazione dopo un po' se prima di una domanda arriva una frase con delle scuse. Rifacciamo il test.

Danny decide di chiedere il significato di cento parole banali in lingua inglese. La macchina risponde sempre "Non capisco". Alla fine ci dice: "Ho parlato con una macchina! Chiedo il significato di frasi diverse e mi arriva sempre la stessa risposta. Inoltre, una persona di una certa cultura qualche parola di inglese la conosce". Il programmatore duplica la banca dati aggiungendo inglese, spagnolo, cinese e tutte le altre lingue della Terra. Rifacciamo il test.

Danny decide di chiedere il significato di una parola in tutte le lingue del mondo. La macchina risponde sempre correttamente. Alla fine ci dice: "Ho parlato con una macchina! Le ho detto una parola in 40 lingue diverse e le conosceva tutte. Nessun essere umano può avere simili conoscenze". Il programmatore toglie tutte le lingue e lascia solo inglese, spagnolo e cinese. Rifacciamo il test.

Danny effettua tutte le prove di prima e alla fine, esausto, ci dice: "Il mio interlocutore risponde a tono, a volte si stufa e mi manda al diavolo, usa vocaboli diversi, conosce solo tre lingue, a volte sta in silenzio e riprende a parlare se mi scuso per averlo trattato male. A me sembra un essere umano".

Abbiamo creato davvero una **Intelligenza Artificiale**?

La questione adesso si sposta su un piano diverso. Il nostro software, che gira su una potente CPU, ascolta le domande tramite un microfono e risponde tramite altoparlanti, **comprende** davvero di cosa parliamo? La **consapevolezza** fa parte del programma? Esiste una procedura che rende la macchina **consapevole** di ciò che ci diciamo?

Analizziamo bene che cosa abbiamo fatto.

Ogni volta che Danny scopriva di aver parlato con una macchina, il programmatore modificava il software, aggiungendo un sistema per evitare quel modo di essere scoperto. La macchina non si correggeva da sola. Come avrebbe potuto farlo? Danny avrebbe potuto solo pensare "*questa è una macchina*" e lei non avrebbe mai saputo di essere stata scoperta. Invece il programmatore veniva informato del problema, interveniva, e quel problema non si verificava più. Ci sarà sempre un modo per scoprire di star parlando con un programma, oppure, prima o poi, i programmatori avranno previsto ed eliminato tutti i modi per scoprirlo?

Sir **Roger Penrose** è un matematico, fisico e cosmologo britannico, famoso per il suo contributo alla "*matematica che spiega la fisica*" e per i suoi contributi alla cosmologia. Laureato all'Università di Cambridge, è professore emerito all'Istituto di matematica dell'Università di Oxford e nel 1988 ha ricevuto, assieme a Stephen Hawking, il Premio Wolf per la fisica.



Daniele Missiroli

Sono nato a Ravenna e vivo a Bologna, dove lavoro come analista programmatore e consulente privacy nell'azienda che ho fondato. Laureato in fisica con lode, sono attratto da numeri, astronomia e scienze in generale, oltre che da fantascienza, fumetti e film. Scrivo libri di fantascienza e ho creato una serie ambientata su un lontano pianeta. Ho scritto anche un libro di matematica sul calcolo mentale.

Roger Penrose ha scritto vari libri divulgativi. In "**La mente nuova dell'imperatore**" e "**Ombre della mente**", descrive la situazione attuale della fisica e affronta i limiti teorici dell'intelligenza artificiale. Egli sostiene che esistono delle differenze intrinseche e ineliminabili fra un software e l'intelligenza dell'uomo. L'uomo, al contrario di un programma, può compiere operazioni che non sono riconducibili alla logica, come sapere la verità di asserzioni non dimostrabili o risolvere il problema della terminazione. Questo problema si domanda se sia sempre possibile, descritta una formula e un determinato input, stabilire se il programma che la contiene terminerà o continuerà la sua esecuzione all'infinito. È stato dimostrato che non può esistere un programma generale in grado di risolvere il problema per tutti i possibili input. La versione più nota del problema è quella proposta nel 1936 dal matematico **Alan Turing**, insieme alla dimostrazione della sua indecidibilità.

È interessante notare che il test per decidere se una macchina può essere scambiata per un essere umano è stato inventato da Alan Turing (il famoso test di Turing del 1950, in cui il matematico si chiede se una macchina può "pensare") e il problema che dimostra l'impossibilità di costruire un'intelligenza identica a quella umana è sempre dello stesso scienziato. Per macchina intelligente, Turing ne intende una in grado di *pensare*, ossia capace di concatenare idee e di esprimerle. Per Turing, quindi, tutto si limita alla produzione di espressioni non prive di significato. In uno dei suoi articoli ha scritto: *«Secondo la forma più estrema di questa domanda, il solo modo per cui si potrebbe essere sicuri che una macchina pensa è quello di essere la macchina stessa e sentire se si stesse pensando. [...] Allo stesso modo, la sola via per sapere che un uomo pensa è quello di essere quell'uomo in particolare. [...] Invece di discutere in continuazione su questo punto, è normale attenersi alla educata convenzione che ognuno pensi.»*

Il filosofo **John Searle** ha proposto una modifica al test di Turing, che ha preso il nome di **Stanza Cinese**, dimostrando l'inattendibilità del test di Turing come prova sufficiente a dimostrare che un software sia dotato di vera intelligenza. Oggi esistono macchine come quella del nostro esercizio che superano facilmente il test di Turing, ma esistono anche milioni di persone che non lo superano.

Nel suo esempio, chi fa il test è un cinese che scrive le domande in cinese. La nostra macchina capisce il cinese e, alla fine del test, quell'uomo sarà convinto di aver parlato con una vera persona. A questo punto, chi sostiene che sia possibile creare un programma che simula compiutamente l'intelligenza umana, dirà che quella IA, oltre a rispondere correttamente, ha anche capito ciò che ha detto, lo ha compreso, quindi è una macchina in grado di pensare.

Ora facciamo lo stesso esperimento sostituendo Danny alla macchina. Lui non capisce una parola di cinese, ma ha a disposizione il listato del programma usato dalla IA, che è scritto in un



Daniele Missiroli

Sono nato a Ravenna e vivo a Bologna, dove lavoro come analista programmatore e consulente privacy nell'azienda che ho fondato. Laureato in fisica con lode, sono attratto da numeri, astronomia e scienze in generale, oltre che da fantascienza, fumetti e film. Scrivo libri di fantascienza e ho creato una serie ambientata su un lontano pianeta. Ho scritto anche un libro di matematica sul calcolo mentale.

linguaggio che lui capisce benissimo, essendo un programmatore. Anche se non comprende il significato dei simboli cinesi, quando vede che a una determinata sequenza il programma risponderebbe in un certo modo, lo fa anche lui e la persona che pone le domande in cinese si trova a ripetere esattamente l'esperimento precedente. Stesse domande e stesse risposte, tutte corrette. Ma questa volta è stato Danny a usare il programma e, pur avendo risposto bene, come ha fatto prima la macchina, non ha capito niente. Non sa se le domande del cinese erano sul tempo che farà domani, sulla sua salute o erano domande di astronomia.

Come può essere, conclude Searle, che due soggetti abbiano eseguito la stessa operazione, ma l'uomo non abbia compreso nulla, mentre la macchina abbia capito quello di cui si è parlato? In termini linguistici:

*Non si può mai ottenere **semantica** (significato) dalla **sintassi** (manipolazione di simboli formali).*

Il grosso problema che ci troveremo ad affrontare in futuro sarà che macchine come quella che abbiamo creato nel nostro esercizio logico saranno scambiate per esseri umani molte meno volte di quanto degli umani saranno scambiati per delle macchine.

Non tratteremo macchine come esseri umani, bensì tratteremo persone come macchine.

Storiella finale.

«Sai» dice un bambino a un amichetto «ho saputo che hanno inventato un computer che fa tutto quello che fa l'uomo, ma lo fa meglio ed è più veloce».

«Lo so» risponde l'altro «mia madre ha chiesto il divorzio e se l'è comprato!»