

# COMPUTER E INTELLIGENZA

ALAN M. TURING

## IL GIOCO DELL'IMITAZIONE

Propongo la domanda: Possono pensare le macchine? Per cominciare bisognerebbe definire 'macchine' e 'pensare'. Le definizioni potrebbero essere formulate sulla base del linguaggio comune, ma sarebbe pericoloso. Infatti andando a cercare come queste siano usate nel linguaggio comune si finirebbe con dare un senso statistico alla domanda Possono pensare le macchine? Neanche fossero le elezioni presidenziali. Ma sarebbe assurdo. Invece di tentare simili definizioni ho pensato di rimpiazzare questa domanda con un'altra simile ma espressa in un linguaggio meno ambiguo.

La nuova forma del problema può essere descritta come un gioco che chiamerò il "gioco dell'imitazione". Partecipano in tre: un uomo (A) una donna (B) ed uno che fa le domande (C).

L'interrogante sta in una stanza diversa dagli altri due. Il gioco per lui consiste nel capire chi è l'uomo e chi la donna. Lui li conosce come X e Y ed alla fine dovrà associare queste due con A e B.

L'interrogante può fare domande del tipo:

C: Può X dirmi la lunghezza dei suoi capelli?

Supponiamo che X sia A, quindi A deve rispondere. È obiettivo di A fuorviare C. La sua risposta potrebbe allora essere:

'I miei capelli sono alla *garçonne* e la ciocca più lunga è di circa nove pollici.

La risposta ovviamente dovrebbe arrivare da un intermediario o tramite scritto. Il terzo giocatore, B, in questo esempio, ha invece il compito di aiutare l'interrogante. Deve dare risposte vere, quindi. Può anche dire 'Sono io la donna, non ascoltarlo!' Ma l'uomo potrebbe dire lo stesso e allora non servirebbe a niente.

Adesso ci chiediamo: 'cosa accadrebbe se il computer prendesse le parti di A nel gioco?'

**Sbaglierebbe l'interrogante altrettanto spesso con il computer al posto di A che con due uomini invece nell'altra stanza? Questa domanda rimpiazza l'originale: "Possono pensare le macchine?"**

### **CRITICA DEL NUOVO PROBLEMA**

Non appena ci chiediamo quale potrebbe essere la risposta a questa nuova domanda uno potrebbe chiedersi se questa nuova domanda sia veramente degna di essere esplorata. Ma questo lo lasceremo perdere poiché non ci porterebbe a nulla.

**Il nuovo problema ha il vantaggio di tracciare una linea abbastanza chiara tra le capacità fisiche ed intellettuali dell'uomo.**

Nessun ingegnere o chimico dichiara di essere capace di creare un tessuto indistinguibile dalla pelle umana. Anche se un giorno si dovesse riuscire a farlo non potremmo non considerare poco sensato cercare di rendere una macchina pensante più umana rivestendola di questa pelle. Il modo in cui abbiamo posto il problema riflette questo fatto nella condizione di C che non può né vedere né toccare A e B. Altri vantaggi di questo criterio possono essere resi evidenti da alcune domande e risposte campione:

D: Per favore mi scriva un sonetto che abbia come soggetto il Quarto Ponte.

R: Non se ne parla. Non potrei mai scrivere poesia.

D: Somma 34957 e 70764.

R: (Pausa di 30 sec) 105621

D: Gioca a scacchi?

R: sì

D: Ho il re su K1 e nessun altro pezzo. Lei ha il re in K6 e torre in R1. Tocca a lei muovere. Che fa?

R: (dopo 15sec. ) torre in R8. Matto.

Il metodo domanda-risposta sembra adatto ad introdurre tutti i campi del comportamento umano che vogliamo includere. Non vogliamo penalizzare la macchina per la sua impossibilità di competere in creazione di bellezza, né vogliamo far gareggiare un uomo in velocità con un aeroplano. Le condizioni del nostro gioco rendono queste mancanze da entrambe le parti irrilevanti.

A e B possono dire qualunque cosa, anche la più falsa, riguardo alle proprie capacità, l'interrogante non chiederà dimostrazioni pratiche.

Il gioco potrebbe essere accusato di essere troppo svantaggioso per la macchina. Se all'uomo fosse richiesto di essere macchina farebbe una gran magra figura, ad esempio per lentezza e imprecisione in aritmetica.

### **Possono le macchine fare qualcosa che possa essere descritto come pensiero ma che sia molto differente da quello che fanno gli uomini?**

È una obiezione molto forte ma alla fine possiamo dire che se la macchina dovesse riuscire a giocare al *gioco dell'imitazione* in maniera soddisfacente potremmo lasciare perdere questa obiezione.

Si potrebbe anche dire che la strategia migliore di queste macchine per giocare alla meglio al gioco potrebbe essere qualcosa d'altro che l'imitazione del comportamento umano.

Potrebbe essere ma è improbabile.

Comunque non ho alcuna intenzione di investigare la teoria del gioco, la cosa migliore è far sì che vengano fornite risposte che potrebbe essere date anche da un uomo.

## **LE MACCHINE IMPIEGATE NEL GIOCO**

La domanda che ci siamo posti nel primo punto non sarà adeguatamente definita finché non spiegheremo che intendiamo con 'macchina'.

È naturale che noi permettiamo l'applicazione alle nostre macchine di ogni genere di tecnica ingegneristica. Possiamo anche permettere che uno o più ingegneri costruiscano una macchina largamente sperimentale che non possa essere chiaramente descritta da chi la ha costruita.

Inoltre escludiamo dalle macchine gli uomini nati nel modo usuale.

È difficile formulare una definizione che si accordi a queste tre condizioni.

[...]

Il fatto che potrebbe essere possibile realizzare un individuo completo da una singola cellula di un solo essere umano ci porta a concludere che non possiamo permettere una qualsivoglia tecnica per realizzare una macchina.

**L'attuale interesse per le *macchine pensanti* è dovuto all'esistenza di *computer elettronici* o *computer digitali*. Seguendo queste suggestioni noi desideriamo che solo macchine del genere **posano partecipare al gioco**.**

[...]

In conclusione ci stiamo chiedendo **non se esistano** già computer digitali in grado di giocare bene *il gioco dell'imitazione* **ma se è possibile immaginare** macchine di tale genere.

[...]

## COMPUTER DIGITALE E MACCHINE UNIVERSALI

[...]

Data la tabella corrispondente ai possibili stati discreti di una macchina digitale è possibile prevedere quello che farà.

Se dotata di una velocità adeguata, una macchina a stati discreti potrebbe imitare il comportamento di una qualsiasi macchina a stati discreti.

Una macchina a stati discreti potrebbe allora giocare nel ruolo di B e la macchina imitatrice nel ruolo di A, e l'interrogante non sarebbe in grado di distinguerle.

La proprietà dei computer digitali di imitare una qualsiasi macchina a stati discreti è descritta affermando che sono ***macchine universali***.

Infatti una importante conseguenza di una simile capacità di imitazione è che, a parte le considerazioni sulla velocità, non serve progettare una macchina per ogni singolo compito, tutto può essere fatto da un solo computer digitale.

Si potrebbe concludere che tutti i computer digitali sono in un certo senso equivalenti.

La domanda "Possono pensare le macchine", può essere trasformata in "Sono immaginabili computer digitali che possono giocare bene al *gioco dell'imitazione*?"

Possiamo fissare la nostra attenzione su un solo particolare computer C. Dotando il computer di una memoria e di una velocità opportuna e fornendogli un programma adeguato, C potrebbe giocare in modo soddisfacente al *gioco dell'imitazione* se prendesse le parti di A e un essere umano giocasse nel ruolo di B?

[...]

### ARGOMENTAZIONE TEOLOGICA

Il pensiero è proprio dell' uomo perché dotato di un'anima immortale

[...]

Nel tentativo di costruire le macchine pensanti noi non desideriamo essere irriverenti usurpando il potere di Dio di creare anime, non maggiormente di quando procreiamo bambini: piuttosto noi siamo in ogni caso strumenti della Sua volontà che provvede a stabilire dove debbano risiedere le anime che Egli crea.

[...]

### L'OBIEZIONE DELLE "TESTE NELLA SABBIA"

Le conseguenze dell'esistenza di macchine pensanti potrebbero essere spaventose, preferiamo pensare che questo non sia possibile.

Noi preferiamo credere che l'uomo sia in qualche modo superiore al resto del creato ed è meglio se si può mostrare che egli è *necessariamente* superiore, in modo tale che il suo potere possa essere ben saldo.

La popolarità dell'argomentazione teologica è chiaramente connessa con questo modo di sentire.

Probabilmente questo punto di vista ha forti radici tra gli intellettuali, che quindi sono portati a credere che *il pensiero* sia la qualità più elevata degli esseri umani e preferiscono fondare la superiorità dell'Uomo su questo potere.

Non credo che valga la pena stare a discutere su questo [...]

## L'OBIEZIONE MATEMATICA

Esistono un certo numero di risultati della logica matematica che possono essere usati per mostrare che ci sono limiti alle possibilità delle macchine a stati discreti.

Il più conosciuto di tali risultati è noto come teorema di Gödel (1931), esso dimostra che in qualsiasi sistema logico sufficientemente potente si possono formulare proposizioni di cui non si riesce a dare una dimostrazione né di esse, né della loro negazione, all'interno del sistema, derivandone così la possibilità dell'incoerenza dello stesso sistema logico.

Sono stati trovati risultati in un certo senso simili, dovuti a Church (1936), Kleene (1935), Rosser, e Turing (1937).

L'ultimo risultato è il più conveniente da utilizzare dato che si riferisce direttamente alle macchine, dal momento che gli altri possono essere usati relativamente ad argomentazioni secondarie: per esempio il teorema di Gödel è utilizzato per rafforzare l'idea che abbiamo motivi per descrivere i sistemi logici in termini di macchine e le macchine in termini di sistemi logici.

Il risultato in questione si riferisce a un tipo di macchina che è essenzialmente un computer digitale con una capacità infinita. Esso afferma che ci sono certe cose che una tale macchina non può fare. Se gli viene richiesto di dare qualche risposta come nel gioco dell'imitazione, ci saranno alcune questioni alle quali darà una risposta errata, o non riuscirà a dare comunque una risposta a ciò che richiederà molto tempo per una risposta.

Ci sono naturalmente molte di tali questioni, ma quello a cui una particolare macchina non è in grado di rispondere potrebbe rispondere un'altra.

Comunque noi stiamo parlando di domande alle quali una risposta "sì" o "no" è appropriata, piuttosto che di questioni tipo: "Cosa pensi di Picasso?"

[...]

**Il risultato matematico è questo: si suppone che questo teorema provi una incapacità della macchina alla quale l'intelletto umano non è soggetto.**

Una breve risposta a questa obiezione è che sebbene si sia provato che ci sono delle limitazioni alle possibilità di una particolare macchina, si è solamente affermato, senza nessun genere di prova, che tali limitazioni non sono applicabili all'intelletto umano.

[...]

Quelli che sostengono l'obiezione matematica dovrebbero, io credo, a maggior ragione accettare *il gioco dell'imitazione* come una base per la discussione.

Quelli che credono nelle due precedenti obiezioni non sarebbero interessati probabilmente a nessun tipo di criterio.

### **L'ARGOMENTAZIONE DELLA CONSAPEVOLEZZA**

[...]

Qualcuno potrebbe obiettare che le espressioni utilizzate dalla macchina, anche se apparentemente sensate, potrebbero essere prive di contenuto e la macchina potrebbe non essere consapevole di quello che dice, nessuna macchina è in grado di gioire per il suo successo o provare dispiacere per i suoi errori.

**Secondo la forma più estrema di questa opinione, il solo modo per cui si potrebbe essere sicuri che una macchina pensa è quello di *essere* la macchina e di sentire se stessi pensare.**

**[...] Allo stesso modo, secondo questa opinione la sola via per sapere se *un uomo* pensa è quella di essere quell'uomo in particolare. [...] Probabilmente A crederà "A pensa ma B no", mentre B crede "B pensa, ma A no". Invece di discutere in continuazione su questo punto, è normale attenersi alla educata convinzione che ognuno pensi.**

[...]

In breve penso che la maggior parte di quelli che sostengono l'argomento della consapevolezza potrebbero essere persuasi ad abbandonarla piuttosto che essere costretti ad accettare una radicale posizione solipsistica. Questi saranno allora probabilmente costretti ad accettare il nostro test.

Non desidero dare l'impressione di pensare che non ci sia alcun mistero intorno alla coscienza. C'è ad esempio qualcosa di paradossale connesso con qualsiasi tentativo di localizzarla.

Io non credo però che questi problemi debbano essere risolti prima di poter dare una risposta alla domanda che ci siamo posti in queste pagine.

## ARGOMENTAZIONI SU VARIE INCAPACITÀ

[...]

Questo tipo di argomentazioni affermano che la macchina non è ingegnosa, non è bella, non è amichevole, non ha iniziativa, non ha senso dell'humour [...], è incapace di amare, non giosce per un buon gelato alla crema, non ha varietà di comportamento [...]

**Credo che tutte queste affermazioni siano fondate sul *PRINCIPIO DI INDUZIONE SCIENTIFICA***

*Un qualunque uomo ha visto nella sua vita migliaia di macchine. Da quello che osserva su di loro egli perviene ad un certo numero di conclusioni generali.*

Esse sono deformi, ciascuna è progettata per un compito molto limitato, sono inutilizzabili non appena si vorrebbe far loro svolgere un compito di poco diverso da quello per cui sono state progettate., la varietà di comportamento di ciascuna di loro è minima, etc etc.

**Naturalmente ci si affretta a concludere che queste sono le caratteristiche di una qualsiasi macchina.**

Molti di questi limiti sono dovuti alle capacità di memoria molto piccole della maggior parte delle macchine.

[...]

L' applicazione del principio di induzione scientifica avviene in gran misura in modo inconsapevole.

Quando un bambino che si è scottato teme il fuoco e mostra di volerlo evitare, dovrei dire che sta applicando il principio di induzione scientifica? ( Naturalmente potrei descrivere questo comportamento in molti altri modi).

Le attività e le consuetudini degli uomini non sembrano essere materia idonea ad applicare il principio di induzione scientifica.

Per poter ottenere risultati affidabili si deve investigare una zona molto ampia nello spazio e nel tempo, altrimenti finisce che facciamo la parte dei bambini inglesi, che sostengono che è stupido imparare il francese dato che tutti parlano inglese.



[...]

La rivendicazione che la macchina non può commettere errori (diversamente dall'uomo), è una rivendicazione piuttosto strana. Verrebbe da rispondere: “ Sono da disprezzare per questo?”

Tentiamo invece di rispondere cercando di capire che cosa si intende con una tale affermazione. Penso che si intenda dire che l'interrogante sarebbe in grado di distinguere facilmente, a causa di questo, la macchina dall'essere umano.

[...]

Questa critica deriva dalla confusione tra due tipi di errori: errori di funzionamento ed errori di conclusione.

Le macchine “ideali” non possono commettere errori di funzionamento.

[...]

**Errori nelle conclusioni si possono avere soltanto quando viene attribuito qualche significato ai segnali in uscita dalla macchina.** [...].

Quando una proposizione falsa viene rappresentata dalla macchina noi affermiamo che la macchina ha commesso un errore nelle conclusioni.

Non c'è alcun motivo per dire che una macchina non possa commettere errori di questo genere.

La macchina potrebbe ad esempio dare in uscita ripetutamente “ $0 = 1$ ”.

Per fare un esempio meno perverso potrebbe avere qualche metodo per trarre conclusioni utilizzando l' induzione scientifica. Dovremmo aspettarci che un tale metodo la porterebbe a dare in alcuni casi risultati errati.

Alla rivendicazione che la macchina non può essere il soggetto dei suoi stessi pensieri si può rispondere solamente se prima si ammette che la macchina possa avere *qualche* genere di *pensiero*.

**Se per esempio la macchina sta tentando di trovare una soluzione dell'equazione**

$$x^2 - 4x - 11 = 0$$

si potrebbe essere tentati di descrivere l'equazione come parte della macchina e come argomento della macchina stessa in quel momento. La macchina può essere usata per costruire i suoi stessi programmi, o per predire gli effetti di alterazioni della sua stessa struttura. In questo senso la macchina può indubitabilmente essere essa stessa materia dei suoi pensieri. [...]

### L'OBIEZIONE DI LADY LOVELACE

Le più dettagliate informazioni sulla *Macchina Analitica* di Babbage provengono da una memoria di Lady Lovelace (1842). In essa ella afferma: "La macchina analitica non ha nessuna pretesa di dare origine ad alcunché. Essa può fare qualunque cosa noi siamo in grado di ordinarle di eseguire. [...]"

Una variante di tale obiezione è che una macchina non è in grado di fare niente di realmente nuovo

Si potrebbe dire "Non c'è niente di nuovo sotto il sole". [...]

Una modo migliore per esprimere tale obiezione afferma che una macchina non può "coglierci di sorpresa".

Le macchine sono in grado di darci risultati inaspettati con grande frequenza, perché non siamo capaci di fare i calcoli che servirebbero per capire che cosa effettivamente aspettarci da loro, o piuttosto sebbene siamo in grado di fare calcoli li realizziamo in modo affrettato, in forma scalcagnata, rischiando.

[...]

---

#### **N.B.**

**Libera traduzione di alcune parti dell'articolo di Turing. Se qualcuno è curioso di saperne di più può leggere l'articolo di Turing scritto nel 1950, che prende in considerazione altri tipi di obiezioni e soprattutto parla in modo più diffuso delle macchine che dovrebbero giocare al *gioco dell'imitazione*, ed inoltre dedica alcune interessanti pagine alle *LEARNING MACHINES*, alle macchine che apprendono.**