

Il test di Turing rivisitato

Introduzione

Il test di Turing si basa sul presupposto che l'intelligenza umana sia in tutto analoga all'intelligenza artificiale, quindi interamente riconducibile alla capacità di elaborare informazioni sulla base di procedure e criteri predefiniti. In una simile prospettiva, è ovvio aspettarsi che il progresso scientifico, rendendo possibile la costruzione di elaboratori sempre più potenti, permetterà prima o poi all'intelligenza artificiale di eguagliare quella degli uomini. In previsione del raggiungimento di un simile traguardo, il matematico Alan Turing mise a punto un test che, nelle sue intenzioni, avrebbe dovuto permettere di stabilire quando ciò sarebbe avvenuto.

Il test di Turing si svolge così. Una macchina computazionale e un essere umano vengono collocati in due stanze distinte. Dall'esterno, un esaminatore, che ignora l'identità degli interlocutori, rivolge loro un certo numero di domande tramite un terminale, ricevendo le relative risposte. Se dopo un certo tempo, l'esaminatore non riesce a riconoscere la macchina dall'uomo, si conclude che la prima si è comportata esattamente come un essere umano. Significa che non esistono differenze di principio tra le modalità di elaborare l'informazione da parte di un computer e i processi intelligenti posti in atto dagli uomini.

Automatizzare la somministrazione del test

Una obiezione – a mio avviso importante – che si può rivolgere al test di Turing, è che l'ideazione delle domande è lasciata interamente alle capacità, inevitabilmente variabili, dell'esaminatore di turno, così pure la valutazione delle risposte ottenute. Non c'è nulla che faccia da argine alla soggettività individuale. Si può ridurre, almeno in parte, questo inconveniente ponendo più persone al posto di un solo esaminatore, in modo che le carenze del singolo siano compensate dalle capacità degli altri. Ma anche così, nessuno potrebbe dirsi ragionevolmente certo che l'eventuale superamento del test di Turing da parte di un sistema artificiale indichi davvero che l'intelligenza artificiale abbia raggiunto il livello di quella dell'uomo, dimostrando così che non esistono differenze di principio tra queste. Rimarrebbe sempre lecito il dubbio che

altri esaminatori più preparati, con domande diverse, avrebbero potuto indurre risposte capaci di permettere il riconoscimento del sistema artificiale dall'essere umano.

Una possibile soluzione a questo problema sembrerebbe essere quella di *standardizzare* completamente il test, stabilendo una volta per tutte quali debbano essere le domande da utilizzare e con quali criteri valutare le risposte ottenute. In tal modo, l'incidenza della soggettività degli esaminatori verrebbe ridotta al minimo e i risultati del test potrebbero essere considerati veramente *oggettivi*. Si potrebbe addirittura pensare di eliminare del tutto l'intervento di esseri umani nella somministrazione del test inserendo, sia le domande da rivolgere al computer sottoposto al test, che i criteri con cui valutare le risposte, in un sistema basato sull'intelligenza artificiale.

Con questi accorgimenti si potrebbe credere di aver raggiunto il massimo dell'affidabilità, sottraendo il test dalla non uniformità delle facoltà umane.

In realtà, la proposta di standardizzare il test di Turing e, ancor più, quella di affidarne interamente la somministrazione a un sistema artificiale, non migliora affatto l'attendibilità del test, nel senso di renderlo più "oggettivo". Non appena si tenta di procedere in tal senso, infatti, ci si trova davanti a problemi enormi, il cui unico sbocco sembrerebbe essere la messa in discussione del presupposto di partenza su cui poggia il test, ovvero la sostanziale analogia tra l'intelligenza umana e l'intelligenza artificiale.

Per meglio comprendere il significato di queste affermazioni, occorre partire da una semplice considerazione: una volta che fossero definite tutte le domande da utilizzare nella somministrazione del test e stabiliti i criteri con cui valutare le risposte, sarebbe relativamente facile per un bravo programmatore intervenire a livello di software sul computer da sottoporre al test per far sì che le sue risposte risultino indistinguibili da quelle che ci si aspetterebbe da un essere umano.

Si potrebbero ovviamente, ideare nuove domande e nuovi criteri di valutazione, nell'ottica di far emergere nuovamente delle differenze tra l'intelligenza dell'uomo e quella del computer. Ma nulla vieterebbe al nostro programmatore di apportare ulteriori modifiche o integrazioni al software del computer per ripristinare la condizione di indistinguibilità tra uomo e macchina.

E' abbastanza ovvio che un simile rincorrersi tra interventi sulla metodologia del test da somministrare e implementazioni a livello di software del computer sottoposto ad esame non potrebbe mai considerarsi concluso, poiché sono sempre immaginabili ulteriori interventi sia da una parte che dall'altra.

Il problema non cambia nella sostanza se invece di stabilire con precisione quali debbano essere le domande che compongono il test, se ne fornisce un numero molto più grande, tra cui selezionare, di volta in volta, con un'estrazione casuale, quelle da utilizzare. Non cambia neppure se, nel caso di un esaminatore-macchina, invece delle domande vengono fornite delle regole (*algoritmi*) con cui costruire ogni volta nuove domande. Anche qui, un bravo programmatore, sia pur con maggior difficoltà, analizzando le domande o i criteri usati per elaborarle, riuscirebbe a far in modo che non vengano rilevate differenze tra le risposte del computer e quelle dell'essere umano.

L'intervento dell'uomo

La verità è che *non esiste alcun insieme finito di domande (o di regole per la loro costruzione) e di criteri per la valutazione delle risposte* – non importa se utilizzati da esaminatori umani o inseriti all'interno di un sistema artificiale – *capace di fungere da metodologia generale,*

che permetta, nel caso di mancata discriminazione tra uomo e computer, *di affermare con certezza la totale sovrapposibilità tra l'intelligenza artificiale e l'intelligenza dell'uomo*. Non esiste, poiché la capacità discriminatoria del test dipende strettamente – anche se ciò è ben lungi dall'essere riconosciuto – proprio dalla componente di *novità* delle domande: dal suo collocarsi, per almeno qualche aspetto, al di fuori delle conoscenze e delle capacità di elaborazione che caratterizzano il sistema artificiale esaminato.

La competizione tra chi elabora domande sempre più difficili o nuovi criteri di valutazione delle risposte, e tra chi apporta modifiche al programma del computer da sottoporre al test, ci mette tuttavia davanti a un fatto incontrovertibile: in questo processo, di cui non si riesce ragionevolmente a intravedere un punto d'arrivo, gli unici soggetti *attivi*, vale a dire in grado di offrire contributi inediti, sono gli esseri umani, mentre i computer (quello sotto esame ed, eventualmente, quello che funge da esaminatore) si limitano a riproporre, anche se variamente elaborato, ciò che era stato precedentemente inserito al loro interno.

Siamo così giunti al punto-chiave dell'intera questione. Il test di Turing dovrebbe dimostrare che non esistono differenze di principio tra l'intelligenza artificiale e l'intelligenza dell'uomo. Ma, per farlo, non può evitare di ricorrere all'intelligenza umana. Infatti, se il test viene standardizzato, cioè reso conforme ai principi che governano l'intelligenza artificiale, perde ogni efficacia. Su un altro versante, abbiamo che qualsiasi modifica o integrazione al test volta ad accrescere la capacità di discriminazione tra essere umano e computer non può essere ideata e messa a punto per mezzo di procedure automatiche, ma richiede l'intervento di un operatore umano. Lo stesso accade qualora si voglia apportare dei miglioramenti al programma del computer da sottoporre al test: c'è bisogno di un programmatore umano che, presa visione dei limiti del computer (rispetto a un determinato stato di avanzamento del test), apporti le opportune modifiche a livello di software.

Intelligenza e creatività

I computer si mostrano migliaia e, in alcuni casi, milioni di volte più veloci degli esseri umani nelle operazioni di tipo logico- matematico. Ma questa maggiore velocità non si traduce in migliori prestazioni in tutte le attività svolte dall'uomo, come sarebbe ovvio aspettarsi se l'intelligenza nel suo complesso coincidesse con la capacità di elaborare l'informazione. Ci sono attività che, nonostante i tentativi finora compiuti, continuano a mostrarsi incontestabilmente al di fuori della portata dell'intelligenza artificiale. Se facciamo riferimento alle esperienze maturate con i sistemi artificiali, tale limite sembra valere, in particolar modo, per quei compiti che richiedono una certa dose di *creatività* per poter essere portati a termine.

La creatività è una facoltà la cui esistenza viene ampiamente riconosciuta dagli psicologi e dalla maggioranza di coloro che si occupano delle mente umana. Il problema sorge quando si cerca di darne una spiegazione (o anche solo una descrizione) che sia coerente con la teoria dell'informazione e, più in generale, con il modello di spiegazione scientifica basato sulla riconduzione a leggi di carattere generale. Il problema consiste nel fatto che la principale caratteristica di ogni atto creativo è quella di porsi, per almeno qualche aspetto, al di là dell'esistente, vale a dire dei criteri regolativi e dei riferimenti di cui si dispone a un determinato istante.

Purtroppo manca ancora una chiara consapevolezza di quanto l'impossibilità di inserimento in un qualsiasi insieme di principi, di norme o di procedimenti predefiniti renda estranea la creatività rispetto ai principi di funzionamento dei sistemi artificiali. Questa mancata consapevolezza spinge alcuni ricercatori a tentar di simulare le doti creative all'interno di

elaboratori appositamente progettati, credendo di trovarsi sulla buona strada per una riproduzione fedele di quelle umane. Sono state costruite macchine capaci di “dipingere quadri”, di “scrivere versi” o di “comporre musica”, con risultati a volte non privi di un certo interesse. Andando però ad analizzare da vicino i prodotti di una simile “creatività”, si scopre immancabilmente che essi sono ottenuti facendo agire il caso all’interno di un certo numero di vincoli. Tali prodotti, si rivelano perciò sempre relativi al bagaglio di informazioni iniziali e alle regole imposte al sistema. Non potrà mai accadere di imbattersi in risultati veramente originali, come avviene in molti casi con gli esseri umani. Non potrà mai accadere – per fare degli esempi concreti – che vengano proposte soluzioni appropriate a un qualche problema ancora aperto, che vengano suggeriti miglioramenti sostanziali in un qualsiasi settore dell’esistenza pratica degli uomini o, traguardo ancora più estremo, che vengano avanzate ipotesi teoriche riguardanti un qualche fenomeno fisico ancora privo di spiegazioni scientifiche.

Ma questo è proprio ciò che si verifica nell’uomo: di fronte a situazioni problematiche inedite, non superabili mediante le conoscenze e le strategie operative disponibili, la nostra intelligenza perviene talvolta a soluzioni per una strada completamente diversa dalle modalità di procedere dei computer. Le facoltà creative permettono di compiere un vero e proprio *salto logico*, non ripercorribile mediante una serie finita di passaggi elementari, sulla base di concetti e i procedimenti pre-definiti.

La creatività è sfuggente, per lo più incontrollabile dalla volontà e non offre garanzie sulla validità dei risultati a cui giunge. Ma la percentuale dei suoi successi è così largamente superiore rispetto a quanto bisognerebbe aspettarsi se ci si affidasse al puro caso, da lasciare pochi dubbi circa la sua esistenza come componente dell’intelligenza ben distinta dalle capacità logico-razionali che l’uomo condivide con i sistemi artificiali...

Considerazioni finali

Il test di Turing sembra poter funzionare, ma ciò avviene soltanto nella misura in cui esso viene applicato in maniera meccanica e acritica secondo le indicazioni del suo ideatore. Appena si assume una prospettiva più ampia e ci si interroga sui criteri con cui costruire le domande e/o valutare le risposte; appena ci si pone il cruciale interrogativo se un test che dovrebbe dimostrare l’indistinguibilità tra intelligenza artificiale e intelligenza umana possa essere somministrato secondo regole e procedure prestabilite, come avviene in tutti gli elaboratori d’informazione, ci si rende conto che “qualcosa non va” negli stessi presupposti su cui si fonda il test. Si comprende che l’intervento dell’uomo è indispensabile per il funzionamento del test, ma anche che tale necessità non può che indicare una differenza sostanziale tra i due tipi d’intelligenza posti a confronto.

E con questo il test di Turing perde qualsiasi interesse, poiché ciò che si proponeva di dimostrare è da considerare del tutto infondato.