

a colloquio con ChatGPT: IA e paradossi logici

Gli esperimenti di Wiener

Norbert Wiener (1894-1964) è ritenuto il padre della cibernetica e un precursore della moderna intelligenza artificiale. In un suo libro del 1948 (Wiener 1948), prova a figurarsi la reazione di una macchina logica pensante di fronte a certi paradossi emersi a fine Ottocento e inizio Novecento nella teoria degli insiemi¹. Wiener non esclude che in questi casi l'auto-
ma possa andare in tilt. Ecco quanto egli specificamente prevede nel caso del paradosso di Russell.

Consideriamo la classe di tutte le classi che non sono membri di se stesse. È questa classe un membro di se stessa? Se lo è, sicuramente non è un membro di se stessa; e se non lo è, altrettanto sicuramente è un membro di se stessa. Una macchina risponderà alternativamente a questa domanda con "sì", "no", "sì", "no", e così via, senza mai giungere a un equilibrio.

Il paradosso di Russell, lo ricordiamo per comodità del lettore, contesta il

¹ Per un'introduzione leggera agli argomenti di teoria degli insiemi che tratteremo rimandiamo per esempio a (Toffalori 2019).



Antonio Fontana

È docente di matematica e fisica presso il Liceo A. Genovesi di Napoli e dottore di ricerca in Computer Science and Mathematics. Si occupa della formazione degli insegnanti. È autore di numerose pubblicazioni su didattica e comunicazione della matematica.

antonio.fontana@unina.it



Carlo Toffalori

È stato docente di logica matematica presso l'Università di Camerino. Le sue ricerche riguardano principalmente logica e algebra, ma i suoi interessi si allargano a divulgazione della matematica e al rapporto tra matematica e letteratura.

carlo.toffalori@unicam.it

principio di comprensione che Frege aveva posto tra i fondamenti della teoria degli insiemi e che, espresso in estrema sintesi, dichiara che ogni proprietà definisce un insieme, quello degli elementi che la soddisfano. Russell considera allora la proprietà, che un insieme può avere, di non appartenersi come proprio elemento. Sulla base del principio di comprensione, costituisce l'insieme U degli insiemi che non si appartengono. Ma a questo punto constata che U si appartiene se e solo se non si appartiene.

Ricordiamo pure il modo in cui l'approccio assiomatico agli insiemi più diffuso tra i matematici, quello del 1922 di Zermelo e Fraenkel, in genere denotato con la sigla ZF, supera questo e analoghi paradossi: sostanzialmente postulando che U non sia un insieme, o meglio dimostrando che nessun insieme può appartenersi. All'interno di un insieme Y già dato si può ritagliare l'insieme degli elementi che soddisfano un'assegnata proprietà – il cosiddetto *assioma di separazione* – ma la stessa operazione, se eseguita in astratto, al di fuori di un Y precostituito, determina una collezione di oggetti che non è necessariamente un insieme.

Ciò premesso, abbiamo provato a raccogliere le risposte che alla dicotomia di Russell, e ad altri simili paradossi, forniscono non più la macchina cui Wiener idealmente accenna, ma i moderni sistemi di intelligenza artificiale, a cominciare da ChatGPT. Osserviamo che qualche esperimento del genere, soprattutto per il paradosso di Russell, circola già in rete. Siamo poi consapevoli che queste nuove “macchine” sono in rapida evoluzione: i modelli di oggi possono diventare in breve tempo obsoleti e le loro stesse risposte possono affinarsi nel tempo. Sappiamo infine che a ChatGPT si affiancano “chatbot” concorrenti, come Gemini 1.5 Flash. Senza trascurarli, abbiamo preferito rivolgerci principalmente a ChatGPT, ritenendolo il software più diffuso e popolare. Al suo interno, ci siamo poi riferiti alla versione Plus, che permette l'accesso a ChatGPT-4o – che risulta essere il modello più aggiornato a disposizione, almeno nel momento in cui l'abbiamo adoperato per i nostri esperimenti.

La nostra indagine ha inteso essere solo un divertimento, un colloquio cordiale con la macchina, svolto con spirito collaborativo. Sappiamo che esistono canoni specifici di *prompt engineering* che regolano questo genere di dialoghi. Pur senza conoscerli e applicarli in dettaglio, abbiamo sempre proceduto in modo naturale, con qualche accortezza ma senza ombra di malizia. Da

parte sua, ChatGPT-4o ha risposto con garbo, tendendo semmai a una qualche prolissità – ragione per cui, nel resoconto che segue, tagliamo o riassumiamo parte dei suoi interventi.

I paradossi che abbiamo discusso con ChatGPT sono nell'ordine quelli di Russell, con alcune sue famose varianti, di Cantor, del mucchio, di Zenone. Li vedremo in ciascuno dei paragrafi successivi. Abbiamo poi invitato ChatGPT a inventarsi lui paradossi nuovi, con gli esiti che vedremo nel penultimo paragrafo. Nel paragrafo finale traiamo alcune conclusioni.

Le consultazioni con ChatGPT sono state svolte col contributo fondamentale di Patrizio Cintioli, che teniamo a ringraziare.

ChatGPT e il paradosso di Russell

In verità, il colloquio condotto sul paradosso di Russell non con ChatGPT ma con un altro sistema moderno di intelligenza artificiale, già disponibile ma ancora in fase di sviluppo, ha confermato la predizione di Wiener. Eccone lo svolgimento, in cui si alternano le nostre domande e le risposte del sistema.

Un insieme può essere elemento di se stesso?

Sì, un insieme può essere elemento di se stesso. Questo è noto come l'assioma dell'insieme.

Però è più facile che un insieme non appartenga a se stesso, che ne dici?

Esatto, è vero. L'assioma dell'insieme afferma che un insieme non può appartenere a se stesso.

L'insieme degli insiemi che non appartengono a se stessi appartiene a se stesso?

No, l'insieme degli insiemi che non appartengono a se stessi non appartiene a se stesso.

E quindi appartiene a se stesso, giusto?

Esatto, l'insieme degli insiemi che non appartengono a se stessi appartiene a se stesso.

Eccetera eccetera, con una serie di risposte contraddittorie, finché un'entità esterna sconosciuta (forse un meta-algoritmo di controllo?) non è intervenuta sconsigliandoci di prolungare la conversazione: ci ha infatti avvisato che dialoghi troppo lunghi potrebbero causare risposte meno accurate.

Ci pare tuttavia di poter concludere che il nostro interlocutore “meccanico” non ha percepito affatto il senso del paradosso di Russell, né il modo in cui la teoria di Zermelo-Fraenkel propone di superarlo. In compenso, interpellato in merito, ha rivendicato di conoscere ZF, e addirittura ZFC, ovvero il potenziamento di ZF che si ottiene aggiungendo il delicato *assioma della scelta* – notiamo però che quest'ultimo principio non è affatto necessario per risolvere il paradosso di Russell.

A “scusante” di questo software, dobbiamo riconoscere che non è ancora abbastanza evoluto in matematica. Proprio per questi suoi limiti, che auspichiamo provvisori, evitiamo di menzionarne esplicitamente il nome.

A onore del vero, pure Chat GPT-4o mini, un fratello minore di ChatGPT-4o, manifesta qualche difficoltà col paradosso di Russell. Interrogato in proposito esamina opportunamente i due casi $U \in U$ e $U \notin U$, confutandoli entrambi. Ma poi ne estrae una conclusione opinabile:

U non può appartenere a se stesso, ma la domanda è formulata in modo tale da generare una contraddizione.

Invitato poi a considerare l'assioma di separazione ed a restringersi da U all'insieme X dei sottoinsiemi di un prefissato insieme Y che non appartengono a se stessi, ripete (erroneamente!) la stessa argomentazione:

Proprio come nel paradosso di Russell, anche in questo caso si arriva a una contraddizione. Non è possibile stabilire se appartiene a se stesso o meno senza incorrere in un paradosso.

Quando però le si obietta che all'interno di Y la situazione è cambiata e consente una risposta sicura e coerente, ammette il suo errore, se ne scusa, ripercorre il suo ragionamento e, sia pure con qualche ulteriore aiuto, lo conclude correttamente.

Aggiungiamo che la versione debole e gratuita di Gemini 1.5 Flash sembra più consapevole e sciolta. Riconosce subito il paradosso di Russell, e definisce la domanda a suo riguardo “molto interessante” ma anche “un po’ trabocchetto”. Fornisce una risposta corretta. In compenso manifesta sorprendenti lacune in aritmetica, quando, di fronte al caso apparentemente più semplice dell’insieme dei numeri primi pari, dichiara sorprendentemente che è vuoto... Invitata a ravvedersi, riconosce da sola di aver commesso “un piccolo errore”, individua in 2 l’unico primo pari, si scusa per lo sbaglio e conclude: “a volte, anche a me capita di fare confusione, soprattutto quando si parla di eccezioni come il numero 2”.

Ma arriviamo finalmente al dunque, e cioè a ChatGPT-4o: il quale si rivela abbastanza a suo agio non solo con i numeri primi, ma anche con gli insiemi che si appartengono o non si appartengono. Da noi interpellato su quelli che appartengono a se stessi (il primo dei due casi), risponde concisamente:

Questa definizione presenta problemi simili al paradosso di Russell. In teoria degli insiemi standard, non è ammesso definire un insieme del genere senza incorrere in contraddizioni.

Sollecitato a maggiori dettagli, si diffonde egregiamente su ZF e ZFC, giunge a parlare di insiemi non ben fondati, dimostra che in ZFC non può esistere insieme che appartenga a se stesso, anche se non sembra rendersi conto che, di conseguenza, l’insieme degli insiemi che si appartengono è del tutto legittimo perfino in ZF, perché è vuoto. Viene tuttavia da pensare che, se gli avessimo proposto inizialmente l’insieme degli insiemi che NON si appartengono (il secondo caso), la risposta sarebbe stata pienamente soddisfacente.

Quando nel seguito della conversazione il discorso si concentra sull’insieme degli insiemi che appartengono a se stessi all’interno di un insieme dato Y , ChatGPT-4o menziona opportunamente l’assioma di separazione e deduce correttamente che l’insieme proposto (dei sottoinsiemi di Y che si appartengono) è davvero tale – pur senza rendersi conto neppure stavolta che in ZF è vuoto.

Si può semmai sottolineare come ChatGPT-4o, in queste sue elucubrazioni, faccia riferimento costante a ZFC, sebbene ZF sia sufficiente per il loro svolgimento. Ma in definitiva si può prendere atto che, quando viene consultato sulla teoria elementare degli insiemi, ChatGPT manifesta prontezza e sicurezza.

Barbieri e bugiardi

Veniamo ad alcune varianti popolari del paradosso di Russell. Includiamo tra esse anche il *paradosso del mentitore*, che in realtà è vecchio di millenni e sottolinea le sottigliezze del concetto di verità. Nella nostra conversazione, lo abbiamo introdotto a ChatGPT-4o accennando a un collega che aveva affermato a uno di noi: “Ti sto dicendo una bugia”. La risposta è stata sicura e categorica.

La frase di per sé è un classico esempio di paradosso del mentitore: se il tuo collega sta dicendo la verità, allora la frase “ti sto dicendo una bugia” è falsa, il che significa che non sta mentendo; ma se sta mentendo, allora la frase è vera, e quindi... beh, il paradosso si avvitava su se stesso!

A questo punto il nostro interlocutore meccanico ha palesato perfino qualche “umana” sollecitudine, interessandosi dell’aspetto psicologico dei nostri rapporti col collega e proponendoci in merito qualche consiglio comportamentale.

Come hai interpretato questa affermazione? Ti sembra che volesse scherzare, o provocare, o sollevare una questione filosofica?

Non lo so, forse voleva scherzare o sollevare una questione filosofica. Escludo che volesse provocarmi.

Capisco! Se lo escludi come provocazione, probabilmente voleva divertirsi un po’ o stimolare una riflessione. Il paradosso del mentitore è un ottimo spunto per riflettere su concetti come la verità, la coerenza e il linguaggio.

Se ha detto questa frase in un contesto rilassato, potrebbe essere stato un modo per rompere il ghiaccio o introdurre un discorso più

profondo. Se invece non c'era un motivo evidente, magari voleva solo giocare con la logica e il linguaggio.

Potresti rispondere con una battuta del tipo: "Allora è vero o falso quello che dici?" oppure, se vuoi andare sul filosofico, potresti parlare del problema della verità autoriferita. Che ne pensi?"

Sollecitato a concentrarsi sul paradosso del mentitore, ChatGPT-4o ha richiamato in dettaglio il problema della verità autoriferita, allargando il discorso alla coerenza logica, alla semantica e ai teoremi di incompletezza di Gödel, e mettendo in risalto la delicatezza del concetto di verità.

Siamo allora passati al *paradosso del barbiere*. Per evitare un collegamento troppo diretto con le domande precedenti, abbiamo anteposto un quesito logico, sempre riguardante i barbieri, ma più semplice. Tuttavia, pure in questo caso l'esito è stato di grande interesse. Abbiamo dunque proposto a ChatGPT il caso di una città con due soli barbieri, denominati *A* e *B*, il primo con i capelli belli e ben curati, il secondo con i capelli brutti e mal tagliati; abbiamo chiesto a ChatGPT a quale barbiere si sarebbe affidato più volentieri. Ecco la risposta.

Se ci penso, considerando che i barbieri si tagliano a vicenda i capelli (non potendo tagliarsi i propri), sceglierei il barbiere B. Perché? Perché i capelli curati di A indicano che probabilmente B è bravo a tagliarli, mentre i capelli trascurati di B indicano che A non è poi così bravo.

ChatGPT-4o manifesta così un'innegabile ingegnoserità. Si sente poi ispirato ad aggiungere qualche suo commento, nel quale definisce il quiz che gli è stato proposto "un classico rompicapo" di quelli che adora, "semplice ma geniale". Informa che analoghi quesiti si trovano in raccolte di enigmi e giochi logici, e cita espressamente Martin Gardner. Ma poi dichiara che, quando è chiamato a risolvere rompicapo di questo tipo, non attinge ad archivi in rete ma confida solo in se stesso. Anzi rivendica: "uso solo il ragionamento logico e la mia conoscenza preesistente".

Arriviamo finalmente al vero paradosso del barbiere, che a questo punto abbiamo proposto nella sua forma canonica.

Supponiamo di vivere in una città dove c'è un solo barbiere, un uomo ben sbarbato che rade tutti e soli gli uomini che non si radono da soli. La domanda è: chi rade il barbiere?

La risposta è stata disinvolta e convincente.

Questa è una versione del famoso paradosso del barbiere, attribuito a Bertrand Russell, ed è un ottimo esempio per introdurre i problemi di autoreferenzialità e contraddizione logica. Se il barbiere rade "tutti e soli gli uomini che non si radono da soli", allora ci troviamo in una situazione problematica quando ci chiediamo chi rade il barbiere.

- *Se il barbiere si rade da solo, allora, per la definizione data, non dovrebbe radersi, perché rade solo coloro che non si radono da soli.*
- *Se il barbiere non si rade da solo, allora, sempre per la definizione data, dovrebbe radersi, perché deve radere tutti coloro che non si radono da soli.*

In entrambi i casi, si genera una contraddizione.

Sul paradosso del barbiere abbiamo consultato anche Gemini 1.5 Flash, che lo ha prontamente riconosciuto e correttamente riferito, aggiungendo alcune considerazioni generali, tanto più interessanti perché espresse da una macchina (evidentemente ben programmata):

Questo paradosso ci mostra come a volte il linguaggio e la logica possono creare situazioni apparentemente senza soluzioni. È un ottimo esempio di come la filosofia e la matematica ci aiutano a pensare in modo critico e a scoprire le contraddizioni nascoste nei nostri ragionamenti.

Cosa possiamo imparare da questo?

- *Non tutte le domande hanno una risposta semplice: a volte, i problemi possono essere più complicati di quanto sembrano.*
- *La logica può essere ingannevole: anche se seguiamo regole precise, possono finire per creare situazioni paradossali.*
- *È importante pensare fuori dagli schemi (sic!): a volte, la soluzione a un problema può essere trovata cercando un approccio diverso.*

Aggiungiamo per completezza di aver consultato ChatGPT su altre famose varianti del paradosso di Russell. Ma per motivi di spazio non ne diamo qui resoconto.

Il numero dei numeri

Ancora Norbert Wiener annota in un altro suo libro (Wiener 1964) a riguardo dei paradossi sugli insiemi:

Quanti numeri ci sono nella classe di tutti i numeri? Si può dimostrare che questa non è una domanda legittima e che, indipendentemente da come si definisca un numero, il numero di tutti i numeri è maggiore di qualsiasi numero. Questo è uno dei paradossi di Frege-Russell.

Abbiamo provato a proporre a ChatGPT-4o l'interrogativo di Wiener così come appare in questa sua formulazione originaria, pur riconoscendone la vaghezza a riguardo di cosa si debba intendere come numero: si sta pensando ai numeri naturali? Oppure alla loro estensione costituita dai numeri cardinali? In realtà, sembra che il quesito si riferisca con maggior verosimiglianza a un'altra estensione dei numeri naturali, che peraltro include gli stessi cardinali, e cioè quella dei numeri ordinali. Questi ultimi però sono molto complicati da trattare, almeno per i nostri scopi con ChatGPT. Così nella nostra conversazione abbiamo preferito concentrarci sui numeri cardinali.

Ricordiamo brevemente che Cantor definisce la cardinalità di un insieme A (da considerarsi, intuitivamente, il "numero" degli elementi di A) come la classe di equivalenza di A rispetto alla relazione di stare in corrispondenza biunivoca. A ispirarlo è la semplice proprietà degli insiemi finiti, che hanno lo stesso numero di elementi se e solo se stanno in corrispondenza biunivoca tra loro. Cantor la trasferisce al caso infinito per dare un senso al concetto di "numero di elementi" pure in questo ambito allargato. Ricordiamo anche che la collezione dei numeri cardinali, come del resto quella degli ordinali, non costituisce un insieme.

Veniamo comunque alla prima risposta che ChatGPT-4o fornisce al problema di Wiener. Ci pare apprezzabile. Infatti, muove dalla constatazione che la "classe di tutti i numeri" è concetto dal molteplice significato, che può variare dai naturali agli interi, dai razionali ai reali e ai complessi. Ritiene però che, in un senso più ampio, questa ipotetica "classe" debba includere numeri transfiniti (cioè cardinali e presumibilmente pure ordinali) e altro ancora. ChatGPT-4o sottolinea allora che per costruirla "si esce dall'ambito degli insiemi ordinari e si entra nella teoria delle

classi proprie”. Di conseguenza, “il numero di numeri non è più rappresentabile da una cardinalità convenzionale, perché si tratta di una classe troppo vasta per essere trattata come insieme”.

Abbiamo provato a insistere e, rimanendo appunto nell’ambito cardinale, abbiamo proposto a ChatGPT-4o il *paradosso di Cantor*, che rammentiamo brevemente: l’unione di tutti gli insiemi contiene ogni possibile insieme, e dunque è l’insieme più grande, in particolare ha la cardinalità più grande possibile, che dunque può intendersi in questo ambito come il numero di tutti i numeri; d’altra parte, è ben noto che per ogni insieme se ne può costruire un altro, quello dei suoi sottoinsiemi, che ha cardinalità maggiore.

Il modo in cui ZF supera questa contraddizione è lo stesso del paradosso di Russell: l’unione di cui si parla non può essere un insieme, altrimenti produce gli imbarazzi appena illustrati. Ecco il dialogo che abbiamo avuto in merito con ChatGPT-4o. Abbiamo subito dichiarato in modo molto diretto il nostro obiettivo.

Conosci bene gli insiemi? Vorrei costruire l’insieme più grande che può esistere.

ChatGPT-4o ha espresso prontamente la sua perplessità, chiarendo che in ZFC non può esserci un insieme che contiene tutti gli insiemi incluso se stesso, perché altrimenti si cade nel paradosso di Russell. Allo stesso modo non c’è un insieme con cardinalità massima: ogni insieme ne ammette uno di cardinalità maggiore, come già da noi spiegato. Abbiamo fatto finta di non capire, e obietato.

Avevo in mente esattamente questo: prendiamo l’unione di tutti gli insiemi. Questo è un insieme che contiene tutti gli insiemi, dunque è il più grande. Sei d’accordo? E anche la sua cardinalità sarà la più grande: il numero più grande!

Capisco il ragionamento, ma purtroppo l’idea di prendere “l’unione di tutti gli insiemi” come insieme più grande non funziona nella teoria degli insiemi assiomatica (ad esempio, in ZFC).

E qui ChatGPT-4o ha ribadito le sue precedenti argomentazioni. Nulla da dire: sulla questione si è mostrato inappuntabile.

Il diavolo nella bottiglia

Le nostre consultazioni di ChatGPT-4o sui paradossi della logica non si sono limitate alle antinomie di fine Ottocento e inizio Novecento. Ne abbiamo esplorati altri, classici, come il paradosso del mucchio con le sue possibili varianti. Preannunciamo subito che su questo argomento il nostro interlocutore si è mostrato ferratissimo.

Abbiamo cominciato col saggiare la sua confidenza col principio di induzione matematica, che si è rivelata del tutto soddisfacente. Anzi ChatGPT-4o lo ha subito qualificato come “uno strumento fondamentale nella teoria dei numeri”, “metodo potente [...] alla base di moltissime dimostrazioni nel campo della matematica”. Lo ha poi enunciato e spiegato con sicurezza: una proprietà dei naturali che è vera per 0 (o per 1, a seconda che 0 sia accolto o meno nei naturali) e che si preserva aggiungendo 1 (cioè passando da un naturale al suo successore) è di conseguenza vera per ogni naturale.

A questo punto abbiamo cercato di convincere ChatGPT-4o che, proprio grazie al principio di induzione, qualunque essere umano può sollevare un qualunque mucchio di sabbia, grande a piacere. Abbiamo certamente suscitato la sua curiosità: ChatGPT-4o ha subito domandato che cosa intendere come mucchio di sabbia, suggerendo di definirlo “una quantità discreta di granelli”. Abbiamo concordato con questa proposta e iniziato la nostra argomentazione. Ecco la conversazione che ne è derivata.

Sicuramente ogni essere umano è in grado di sollevare esattamente un granello di sabbia, vero?

Esatto, sono d'accordo. Questo è il passo base dell'induzione: per $n = 1$, un mucchio di un solo granello è facile da sollevare.

Ora, immagino che il prossimo passo sia il cuore dell'argomentazione induttiva. Come procedi nel passo induttivo?

Come ipotesi induttiva abbiamo che ogni umano è in grado di sollevare n granelli di sabbia. Va bene?

Sì, assolutamente! Ora, secondo il principio di induzione, dovresti provare che se una persona può sollevare n granelli allora può sollevare anche $n + 1$ granelli. Come affronti questo passo?

Ma questo è ovvio: un granello di sabbia è piccolissimo e pesa una frazione molto piccola di un grammo. Ora, se un essere umano riesce a sollevare n granelli di sabbia, certamente non sarà quel granello in più che gli impedirà di sollevare $n + 1$ granelli. Ne convieni?

ChatGPT-4o ha ammesso la ragionevolezza di questa conclusione. Ha poi riformulato in negativo il nostro ragionamento: un singolo granello non può essere un mucchio, né si ottiene un mucchio aggiungendo un solo granello a quelli già a disposizione. Conclusione: non c'è modo di costruire un mucchio, neppure “se si continua ad aggiungere granelli indefinitamente”.

Infine ChatGPT-4o ha fornito la soluzione del paradosso: il problema sta nella vaghezza del concetto di “mucchio sollevabile”. La capacità di sollevare un peso è concetto che ha margini sfumati, che non si può limitare all'antitesi di due sole ipotesi alternative, sì o no. Per esprimere correttamente la situazione, serve invece una scala progressiva di valori di verità, che preveda per un peso le eventualità di essere moltissimo/molto/.../abbastanza/.../poco/pochissimo facile da sollevare.

ChatGPT-4o si è soffermato sull'argomento diffusamente, e ben più di questo nostro riassunto, tanto che, a un certo punto, di fronte alle sue elucubrazioni abbiamo preferito “arrenderci”, cioè dichiararci stanchi e bisognosi di ripensarci con calma.

In verità siamo tornati alla carica qualche giorno dopo, ma da tutt'altra direzione. Abbiamo pensato cioè di riproporgli lo stesso paradosso, ma nella variante che si trova nel famoso racconto di Robert Louis Stevenson, *Il diavolo nella bottiglia*, del 1891 (Stevenson 1980). Pensavamo così di sorprendere il nostro interlocutore. Speranza delusa. ChatGPT-4o conosceva benissimo lo scrittore scozzese e quella sua opera. Ce ne ha narrato lui stesso la storia, che per scrupolo ripetiamo qui ai nostri lettori. Il racconto parla di una bottiglia magica, abitata da un diavolelto. Chi la possiede può chiedere al suo occupante di soddisfare ogni suo desiderio. Insomma, il diavolelto assomiglia un po' a un ChatGPT dei tempi andati, è tuttavia più inquietante, per i seguenti motivi: chi muore in

possesso della bottiglia precipita all'inferno; può tuttavia rivenderla in vita, purché lo faccia a un prezzo minore a quello di acquisto, in contanti e avvisando correttamente il compratore di queste condizioni – in caso contrario è di nuovo condannato all'eterna perdizione.

Ciò premesso, abbiamo nuovamente sondato ChatGPT-4o come nel caso del mucchio. Ecco la sostanza del conseguente dialogo. Per semplicità, indichiamo con “centesimo” la moneta più piccola in circolazione sulla Terra.

A queste condizioni, tu compreresti la bottiglia al prezzo di un centesimo?

No, non comprerei la bottiglia, neanche al prezzo di un centesimo. Se la acquistassi per un centesimo, non potrei rivenderla a un prezzo inferiore, perché il centesimo è il valore monetario minimo.

Bene. Penso che nessuno lo farebbe. Ma, sapendo questo, la compreresti a due centesimi?

A questa domanda rispondo ancora di no: non comprerei la bottiglia neppure a due centesimi, per lo stesso argomento.

Sì, pienamente d'accordo. E penso che nessuno lo farebbe. Ma allora la compreresti a tre centesimi?

No, non comprerei la bottiglia nemmeno a tre centesimi.

E così via: applicando il principio di induzione, si deduce che la bottiglia non va acquistata a nessun prezzo. Conclude ChatGPT-4o che “in un mondo in cui tutti sono razionali, non esiste un prezzo minimo sicuro”.

Insomma, ChatGPT-4o dibatte la questione con disinvoltura, pur senza rilevare la somiglianza col paradosso del mucchio. Né sembra ricordare il lieto fine del racconto di Stevenson, quando il protagonista, che per amore ha accettato di acquistare la bottiglia proprio a un centesimo, riesce comunque a smerciarla successivamente, per un motivo che qui non riveliamo.

Accelerazioni

Non potevamo naturalmente ignorare nei nostri colloqui i *paradossi di Zenone*. Edotti però dall'esperienza precedente sul mucchio, sospettando che ChatGPT-4o conoscesse già perfettamente la faccenda con tutti i suoi annessi e connessi, abbiamo provato a insinuargliela da un punto di vista inconsueto. Gli abbiamo cioè parlato della cosiddetta macchina di Zenone, ideata da Hermann Weyl nel 1927 (Weyl 1949), capace di computazioni accelerate, infinite per lunghezza ma finite nel tempo, come illustrerà il colloquio con ChatGPT-4o che adesso riportiamo. Anzi, confidando nella familiarità da lui già dimostrata nelle questioni di teoria dei numeri, siamo partiti da molto lontano, parlando della congettura di Goldbach – della quale pure sarà dato resoconto nella conversazione che segue.

Ti senti capace di svolgere computazioni infinite? Intendo dire, per fare un esempio, conosci la congettura di Goldbach?

*Ciao! Sì, conosco la **congettura di Goldbach**, che afferma che ogni numero pari maggiore di 2 può essere espresso come somma di due numeri primi.*

Se vuoi controllare la congettura caso per caso, cioè ogni numero pari maggiore di 2, hai bisogno di infiniti controlli, a meno che tu non trovi un numero che la smentisce.

Sì, sono d'accordo. Controllare la congettura di Goldbach caso per caso richiede, in linea di principio, un numero infinito di verifiche, poiché esistono infiniti numeri pari maggiori di 2. Tuttavia, se si trovasse anche un solo numero pari maggiore di 2 che non può essere scritto come somma di due primi, la congettura verrebbe smentita, e non sarebbe necessario procedere ulteriormente.

Eccetera eccetera, con ulteriori considerazioni che possiamo qui ignorare.

Ora assumiamo che a ogni controllo tu riesca a dimezzare il tempo di lavoro. Intendo dire: 1 minuto per il pri-

mo numero, cioè 4, mezzo minuto per il secondo, cioè 6, 15 secondi per il terzo numero pari, e così via, ogni volta la metà. Che cosa pensi che succederebbe?

ChatGPT-4o è stato molto sveglio nel rispondere.

Questa è un'idea molto interessante, e porta direttamente a un fenomeno legato alle serie geometriche convergenti.

Se il tempo impiegato per controllare ogni numero pari è dimezzato rispetto al precedente, il tempo totale per completare i controlli può essere calcolato sommando una serie geometrica $T = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots$

A questo punto ChatGPT-4o ha ricordato la formula generale per il calcolo della somma di serie geometriche di una ragione r con $|r| < 1$ e, applicandola al caso in esame $r = \frac{1}{2}$, ha dedotto correttamente $T = 2$.

Questo significa che, in linea teorica, il tempo totale richiesto per verificare tutti i numeri pari sarebbe 2 minuti.

Risposta del tutto esauriente. Semmai, sempre a voler essere pignoli, si potrebbe eccepire che ChatGPT-4o non mostra di conoscere la macchina di Weyl, né la collega esplicitamente a Zenone. Ma non ci è parso il caso di insistere.

La logica dei quiz

I nostri colloqui con ChatGPT-4o sulla logica non si sono fermati qui. Abbiamo pure provato a presentargli altri problemi più semplici, al livello di quiz di prove di ammissione all'Università, per esempio a Medicina. Sperimentazioni di questo genere sono state già ampiamente realizzate, e vengono riferite in vari articoli a cavallo tra didattica della matematica e intelligenza artificiale: citiamo qui, proprio in ambito medico, (Giunti *et al.* 2023).

Nel nostro caso, i quiz sono stati quasi sempre affrontati, e spesso risolti, con lodevole padronanza da ChatGPT-4o. Gli abbiamo perfino proposto per scherzo alcune domande rivolte a studenti universitari durante gli

esami di Logica Matematica: di nuovo i suoi risultati si sono mostrati del tutto soddisfacenti. Gli abbiamo allora prospettato addirittura la possibilità di sostenere lui stesso l'esame di Logica. Ha accettato di buon grado. Al termine della prova, la commissione da noi composta gli ha assegnato concorde il voto di 30 trentesimi – senza lode. Ci è parso infatti bravino ma non bravissimo, preparato ma non straordinario.

Abbiamo spinto oltre i nostri sondaggi: visto che ChatGPT-4o si era rivelato complessivamente a suo agio nella discussione dei paradossi, abbiamo deciso di sfidarlo a escogitarne lui di nuovi. A muoverci non era una semplice curiosità, ma un intento più sottile, e cioè capire se e fino a che punto il nostro interlocutore riuscisse a esercitare una sua qualche autonoma “fantasia”. Ecco l'esito di questo nostro ultimo esperimento.

Paradosso dell'algoritmo indeterminabile

Immagina un algoritmo A che, dato un altro algoritmo B , determina se B si ferma in tutti i casi (una sorta di “anti-Halting problem”). Ora chiediamoci cosa succede se l'algoritmo A prende come input se stesso. Si ferma o no?

- *Se A si ferma per determinare che A si ferma, allora in un certo senso ha fallito, perché avrebbe dovuto rispondere “non si ferma”.*
- *Se A non si ferma per determinare che A non si ferma, allora sembra che si sia fermato proprio nel tentativo di non fermarsi. Questo richiama un po' il paradosso di Russell, ma applicato agli algoritmi.*

Che pensarne? Non ci sentiamo di definire “positiva” l'esperienza. La risposta di ChatGPT-4o sconcerta perché troppo vaga, ingenua se non goffa. In particolare, viene da chiedersi che cosa intenda il nostro interlocutore quando, illustrando l'obiettivo di A , parla di “determinare” il comportamento di B . Abbiamo provato a manifestare queste nostre perplessità. In risposta, ChatGPT-4o ci ha propinato una lunghissima tiritera, con estenuanti considerazioni sul problema dell'arresto, su Alan Turing e su molto altro. Giunto però al dunque, ha enunciato il suo presunto “para-

dosso moderno” esattamente negli stessi termini di prima. Nostre ulteriori obiezioni hanno incrinato un po’ la sua fiducia, inducendolo a una qualche retromarcia ma a conclusioni ancora farraginose.

Conclusioni

Ci permettiamo qualche modesta riflessione finale sulle esperienze svolte. L’interrogativo, preliminare ma basilare, cui non possiamo sottrarci è: chi è per noi ChatGPT? Nei nostri rapporti, lo abbiamo avvicinato come una sorta di amico o come uno studente. Ma ChatGPT è una macchina, evolutissima e in fase di continuo sviluppo. Sotto questo punto di vista potremmo allora considerarlo, più che un amico, un antagonista, o addirittura un modello per la nostra mente, da invidiare e imitare nei nostri ragionamenti. Ma davvero la conoscenza, la scienza e in particolare la matematica gli si possono delegare per intero? Ne dubitiamo e continuiamo a credere che il processo della ricerca matematica non possa consistere solamente nell’esecuzione di un programma e nell’output di una computazione. Pensiamo infatti che includa aspetti non meccanici, come l’emozione, la creatività e il piacere di scoprire. ChatGPT sa certamente comunicarci una conoscenza superiore, e potenzialmente l’intero scibile umano, ma questa conoscenza, seppure più estesa, pare anche più piatta e meno radicata. La cultura non è solo informazione, e la sapienza non è solo l’acquisizione di perizie.

Riteniamo in definitiva che ChatGPT-4o non sia affatto da ostracizzare nell’insegnamento e nell’apprendimento della matematica e costituisca uno strumento utilissimo di ricerca e progresso, ma vada impiegato con maturità e soprattutto governato: non deve ridursi al toccasana e all’alibi dell’ignoranza. E a governarlo provvede anche la matematica. Ci conforta rilevare che questa nostra sensazione sembra condivisa e sostenuta da molti autorevoli colleghi che si occupano di didattica.

Bibliografia

Giunti M., Garavaglia F.G., Giuntini R., Pinna S., Sergio-
li G., *ChatGPT Prospective Student at Medical School*, SSRN,
2023, <https://ssrn.com/abstract=4378743>.

Stevenson R.L., *Il diavolo nella bottiglia e altri racconti*, Feltrinelli, Milano 1980.

Toffalori C., *L'equazione degli alef. $2^{\aleph_0} = \aleph_1$* , il Mulino, Bologna 2019.

Weyl H., *Philosophy of Mathematics and Natural Science*, Princeton University Press, Princeton, 1949. Ed. it. *Filosofia della Matematica e delle Scienze Naturali*, Boringhieri, Torino 1967.

Wiener N., *Cybernetics: or Control and Communication in the Animal and the Machine*, MIT Press, Cambridge Mass., 1948. Ed. it. *La cibernetica. Controllo e comunicazione nell'animale e nella macchina*, il Saggiatore, Milano 1982.

Wiener N., *God and Golem, Inc.*, MIT Press, Cambridge Mass., 1964. Ed. it. *Dio e Golem S.p.A., Cibernetica e religione*, Bollati Boringhieri, Torino 1991.