

FISICA PER TUTTI

I primi passi termodinamici di ChatGPT

Stefano Oss

Dipartimento di Fisica, Università di Trento, Trento, Italia

Riassunto. In questo lavoro vengono esposti e brevemente analizzati degli esempi di uso esplicito di ChatGPT, uno dei più famosi e discussi assistenti basati linguaggi estesi di intelligenza artificiale, nell'ambito della termodinamica classica. Si affrontano dei casi di difficoltà varia e di interesse diretto nelle applicazioni della didattica, ovvero con espliciti riferimenti alle possibilità di dialogo e tutoraggio che l'assistente virtuale è in grado di sostenere in momenti di studio e apprendimento. Anche se non vengono proposte valutazioni quantitative e comparative, vista l'enorme variabilità e la giovanissima età di questi linguaggi, l'esplorazione qui riportata evidenzia molte difficoltà nel condurre a termine compiti anche non eccessivamente complessi. Il passaggio tra "la teoria e la pratica", tanto impegnativo nell'apprendimento delle scienze fisiche, con un particolare accento proprio nel campo della termodinamica, non verrebbe facilitato —o addirittura ostacolato o quantomeno reso inutilmente più complesso— per il tramite di ChatGPT o suoi consimili.

1. Introduzione

Ci si può chiedere quali siano i rischi maggiori che si corrono nell'utilizzare gli "assistenti" basati sull'uso di intelligenza artificiale, come ChatGPT (Generative Pre-trained Transformer, o Trasformatore Generativo Pre-addestrato), nelle attività di didattica, in generale, e qui in particolare delle scienze fisiche, connesse ai processi di apprendimento e insegnamento. Un possibile riassunto potrebbe essere il seguente, per rompere il ghiaccio con una lista di punti critici e senza girare troppo attorno al problema:

- 1) *Mancanza di comprensione profonda:* ChatGPT è un modello di linguaggio basato su dati testuali e potrebbe non avere una comprensione completa dei concetti fisici. Gli studenti potrebbero ottenere risposte accurate in termini di terminologia, ma l'assistente potrebbe mancare di una vera comprensione dei concetti sottostanti. Questo potrebbe portare a una superficialità nell'apprendimento.
- 2) *Dipendenza dalla tecnologia:* L'uso eccessivo di assistenti virtuali potrebbe portare a una dipendenza dalla tecnologia e ridurre l'interazione studente-insegnante. Gli studenti potrebbero perdere l'opportunità di sviluppare competenze critiche come il pensiero critico e la risoluzione dei problemi se si affidano troppo a tali assistenti.

- 3) *Mancanza di personalizzazione*: ChatGPT può fornire risposte standardizzate e non personalizzate, il che potrebbe non soddisfare le esigenze individuali degli studenti. Gli insegnanti sono in grado di adattare il loro insegnamento in base alle esigenze specifiche degli studenti, mentre un assistente virtuale potrebbe non essere altrettanto flessibile.
- 4) *Rischi di frode e copiatura*: Gli studenti potrebbero utilizzare ChatGPT per ottenere risposte immediate ai compiti o alle domande degli esami senza sforzo personale. Ciò potrebbe incoraggiare la copiatura e la frode accademica.
- 5) *Mancanza di feedback qualitativo*: Gli insegnanti forniscono non solo risposte alle domande degli studenti, ma anche feedback qualitativo e orientamento. L'uso esclusivo di assistenti virtuali potrebbe privare gli studenti di questo feedback prezioso.
- 6) *Privacy e sicurezza*: L'uso di assistenti virtuali potrebbe comportare la condivisione di dati personali o sensibili degli studenti, sollevando preoccupazioni sulla privacy e sulla sicurezza dei dati.

In altre parole, emergerebbero da subito delle criticità molto pronunciate che devono mettere in guardia gli operatori didattici, a ogni livello, sulle possibili ricadute negative dell'utilizzo di ChatGPT o altri "bot" (che sta per "robot", ossia sistema o macchina di supporto e automazione) senza un attento e sistematico lavoro di supervisione e di personalizzazione dell'insegnamento. Se a questo fatto, per alcuni molto ovvio, per altri forse molto sottovalutato almeno in questo periodo, aggiungiamo quello che l'elenco sopra esposto è stato prodotto proprio da ChatGPT "in persona", è chiaro che la situazione è decisamente preoccupante. Se non altro a questo assistente non sembra mancare una buona dose di autocritica.

Anche se questo lavoro non prevede una rassegna delle pubblicazioni e delle ricerche nel settore dell'utilizzo dell'IA nella didattica della fisica (per la quale si può per esempio fare riferimento a [1]), lo scenario attuale (nazionale in primis, ma anche a livello internazionale [2–9] le cose non vanno molto diversamente) è che ci siano più chiacchiere che reali e affidabili sperimentazioni. La letteratura è molto sparsa e per ora limitata a relazioni su sperimentazioni non strutturate e pubblicate su archivi temporanei [10–14] e non su riviste di settore (che, peraltro, sono non meglio definite o comunque a cavallo tra ambiti informatici, linguistici, socio-psicologici, cognitivi, didattici). A ciò si aggiunga che le dicerie e gli utilizzi molto occasionali dell'intelligenza artificiale nei progetti di affiancamento a percorsi didattici abbondano, anche e forse soprattutto per causa del solito fenomeno di diluizione delle conoscenze e della diffusione di opinioni non verificate né verificabili veicolate dai frequentemente devastanti *social*. Mancano in ogni caso dati stabili e statistiche robuste che aiutino a farsi un quadro attendibile. Del resto ChatGPT (ma anche altri assistenti, come Google Bard/Gemini [15] e Microsoft Copilot [16], e non sono gli unici [17]) sono entrati in campo solamente negli ultimi mesi, anche se, in realtà, i motori di questi "bot" erano attivi da molto più tempo in settori di enorme importanza operativa per la nostra società, come nella finanza, nella medicina, nelle comunicazioni, nei trasporti e così via [18]. Le interfacce con linguaggio naturale, o addirittura vocale, sono invece balzati in evidenza da pochissimo tempo e hanno invaso pressoché ogni ambito sociale,

sia ludico che artistico, educativo e narrativo e molto altro. Sono un “giocattolo” molto nuovo e non bene conosciuto, neppure dai loro stessi creatori, per certi versi. Proprio per questo motivo è essenziale essere estremamente cauti prima di adottare questi “grandi modelli linguistici” o LLM come efficienti ed efficaci assistenti per fare scelte importanti, soprattutto nel campo della didattica.

In questo lavoro si è scelto di non procedere con una sperimentazione diretta rivolta alle utenze primarie della scuola, ovvero studenti e studentesse, proprio perché si ritiene per ora molto più urgente e illuminante approfondire qualche aspetto operativo, procedurale e contenutistico di uno di questi assistenti, ChatGPT. Per puntare fin dall’inizio a un motore abbastanza stabile e affidabile, anche se non al massimo delle capacità che l’hardware di oggi può consentire, si è deciso di adottare l’interfaccia 4.0 (a pagamento, circa 20 € mensili) e, il più spesso possibile, di confrontare i risultati ottenuti con la versione 3.5 (gratuita). Tra i punti di forza della versione 4.0 sulla 3.5 si cita quello di consentire lo scambio di file di vario formato (pdf, grafici, tabelle, etc.) e di interagire più in profondità discorsiva con l’interlocutore. Anche il database da cui attinge per informazioni bibliografiche e altre fonti è più aggiornato.

Un’analisi dettagliata delle modalità con le quali ChatGPT potrebbe affiancarsi a studenti e docenti nel lavoro di studio e apprendimento delle scienze fisiche è molto complesso e si può ben immaginare che potrà richiedere tempi molto lunghi. Il problema è, infatti, che l’ampiezza di questa disciplina e le differenti difficoltà concettuali e applicative dei moltissimi campi nei quali spaziare —pur sempre limitandosi a livelli e schemi tipici della didattica— sono tali da rendere un compito arduo giungere a conclusioni generali. L’impressione che si ha durante l’utilizzo di questi bot è che, in funzione delle modalità con le quali domande e altre richieste vengono poste dall’utente, come pure al variare dello specifico settore disciplinare, si ottengano risposte con uno spettro di affidabilità, correttezza, spendibilità estremamente variabile.

Per evitare, in questa fase preliminare di esplorazione, un’eccessiva dispersione su una grande vastità di temi come quella delle scienze fisiche, si è deciso di considerare un insieme molto limitato di casi di studio facenti capo alla termodinamica classica, con specifica attenzione all’applicazione dei suoi principi fondanti a qualche problema non eccessivamente complesso ma neppure da essere considerato troppo banale in contesto didattico. Può capitare che, per motivi legati alle poche ore disponibili, la termodinamica venga “accodata” in appendice dei corsi di fisica generale universitaria o direttamente trasferita a quelli più ingegneristici di fisica tecnica. Nei licei viene trattata ovviamente come meglio si può —ci sono varie linee guida e resoconti di ricerca relative a sperimentazioni di notevole interesse [19,20]— ma non è sempre considerata degna di pari attenzione delle parti dedicate alla meccanica. I motivi sono vari ma, semplificando e riducendo al massimo un problema piuttosto articolato, spesso questa disciplina evidenzia difficoltà concettuali che si trasferiscono inevitabilmente su altrettante critiche situazioni nell’impostazione e svolgimento di esercizi e problemi.

In questo lavoro abbiamo rivolto l’attenzione alle modalità di interazione con ChatGPT e con le risposte che questo linguaggio fornisce. Va subito osservato che un modo tipico di esprimersi di questo assistente è abbastanza prolisso e ripetitivo e si nota una propensione ad anticipare alla risposta vera e propria svariate considerazioni

più generali e teoriche, quasi si intendesse forzare uno stile didattico e maieutico durante l'interlocuzione. Ciò comporta, in ogni caso, risposte piuttosto estese e, in questo lavoro, si è deciso di tagliare le parti più lunghe per concentrarsi sugli aspetti sostanziali dei “dialoghi”. Infine, si è adottata in tutto il lavoro la lingua italiana: questi assistenti non hanno particolari problemi di comprensione e di scrittura in molte lingue tra quelle a più ampia diffusione, un altro fattore che rende particolarmente attraente l'utilizzo di motori di IA di questa specie, in grado infatti di essere universali, letteralmente e linguisticamente senza confini. Spesso sono stati fatti confronti con l'equivalente dialogo tenuto in lingua inglese (quella “nativa” di ChatGPT) e non si sono trovate particolari differenze.

Volendo partire subito con una possibile strategia di interazione che permetta di evidenziare punti di forza e di debolezza didattici dell'interlocutore artificiale, muovendosi quindi nell'ambito della termodinamica, è importante immaginare quali potrebbero essere le modalità di accesso dell'utenza “fragile” degli studenti (d'ora in avanti, per non appesantire il testo, si farà utilizzo spesso di nomi declinati al maschile senza ovviamente nessuna mancanza di riguardo per la questione di equilibrio di genere). Ci sono —non solo nelle questioni legate alla termodinamica, peraltro— due principali tipologie di “*prompt*”, ovvero di interrogazione dell'assistente. Nella prima si cerca aiuto/consiglio relativamente all'enunciato corretto di principi e leggi della fisica; nella seconda, invece, l'attenzione è rivolta alla risoluzione di esercizi e problemi. Senza qui entrare nel merito pedagogico e metodologico di questo schema, si intravede però immediatamente che l'utilizzo del linguaggio artificiale che presenta più lati oscuri e rischi per l'utenza fragile suddetta è sicuramente il secondo, con il quale si cercano inevitabilmente vie brevi per giungere alla soluzione del problema assegnato. Per quanto possa sembrare comunque interessante l'opportunità di confrontarsi con un interlocutore instancabile e molto acculturato come il chatbot utilizzato, è molto importante osservare e non sottovalutare il fatto che i dialoghi con esso hanno senso solamente se l'assistente virtuale è esperto e capace di esercitare un'azione maieutica. Vedremo nel seguito che questi caratteri non emergono molto convintamente dall'ambiente IA. Si può però iniziare l'interrogazione del bot parlando di questioni teoriche, ossia di tematiche che non dovrebbero lasciare eccessivo spazio all'ambiguità e al dubbio. In ciò appare che ChatGPT sia in una fase di addestramento avanzato e affidabile in molte situazioni. È abbastanza difficile riuscire a mettere in crisi l'assistente riferendosi a una varietà ampia di argomenti e capitoli della termodinamica, inclusi quelli relativi al trasporto dell'energia termica e alle fasi delle sostanze, tematiche che usualmente non sono affrontati nelle scuole di II grado, ma che sicuramente potrebbero costituire un ottimo argomento di approfondimento e ricco di importanti applicazioni.

Anche la risoluzione di esercizi (che qui distinguiamo dai problemi semplicemente per il fatto che i primi sono dirette applicazioni di relazioni e leggi già assegnate, mentre i secondi richiedono processi anche molto impegnativi e creativi di impostazione e adattamento di vari modelli) è un terreno abbastanza congeniale all'assistente. In questo lavoro, per semplici questioni di spazio, non verranno considerati in dettaglio gli esercizi che sono digeribili dall'IA. Però, in un contesto didattico, è bene qui sot-

tolineare che sono proprio queste tipologie di impegno e di studio, fondamentali per gli studenti, a essere pesantemente modificate e distorte dall'utilizzo degli assistenti virtuali a casa o a scuola senza monitoraggio e verifica attenta di quanto viene prodotto, soprattutto quando l'argomento in questione è nuovo e non c'è dunque modo di auto-verificare la bontà delle soluzioni. Soprattutto, in aggiunta, qui subentra inevitabilmente e minacciosamente la sostituzione dell'impegno e della fatica dell'apprendimento da parte dello studente grazie al sistema automatico che ha svolto il "lavoro sporco". Per quanto sia definito come tale, l'impegnarsi a organizzare esempi, calcoli, operazioni e stime numeriche, è parte integrante della crescita cognitiva, di competenze e abilità, che è componente irrinunciabile di uno sviluppo intellettuale solido, personalizzato: una base fondamentale per consentire di impostare, affrontare e risolvere i problemi più complessi.

Vediamo ora qualche situazione concreta di "interazione termodinamica" con ChatGPT. Vista la natura "non-deterministica" del sequenziamento delle risposte fornite dall'assistente [21], si tenga presente che alle medesime domande poste a distanze anche brevi di tempo si otterranno forme differenti sia sintatticamente che nell'ordine dei passi logici seguiti.

2. I problemi irrisolti

Un primo caso di studio parte dalla domanda relativa alla possibilità che un sistema termodinamico, in un ciclo chiuso, possa assorbire 1600 J di calore dopo aver eseguito un lavoro di 1000 J, interagendo con termostati a temperature di 0 °C, 100 °C e 400 °C. Viene chiesto esplicitamente a ChatGPT di rispondere facendo uso dei principi della termodinamica, includendo anche l'applicazione del teorema di Carnot. Sappiamo infatti che una macchina termica operante con qualsiasi numero di serbatoi di calore non potrà mai avere un rendimento maggiore della macchina ideale —necessariamente di Carnot— che opera tra i due termostati alle temperature estreme del ciclo in esame. Nel caso assegnato, è immediato calcolare l'efficienza di Carnot date le temperature massima e minima di funzionamento:

$$\eta_{\text{Carnot}} = 1 - T_{\text{min}}/T_{\text{max}} = 1 - \frac{(0 + 273) \text{ K}}{(400 + 273) \text{ K}} = 0.59,$$

mentre, a partire dal calore assorbito e dal lavoro prodotto per ciclo, l'efficienza prevista è data da

$$\eta = \frac{W}{Q_{\text{assorbito}}} = \frac{1000 \text{ J}}{1600 \text{ J}} = 0.625,$$

che è maggiore del rendimento massimo di Carnot. Questo viola il secondo principio della termodinamica e dunque conduce subito alla non realizzabilità del ciclo proposto nel testo. ChatGPT prende una strada differente. Enuncia con dovizia di dettagli (e correttamente) il primo e il secondo principio nel formato tradizionale, ovvero parlando di conservazione dell'energia e del predicato di Clausius; esprime poi il teorema

di Clausius con riferimento al limite superiore per l'efficienza di motori termici operanti tra due temperature. Aggiunge subito al primo principio questa considerazione (da qui in avanti, le frasi in italico sono trascrizioni “copia e incolla” delle affermazioni dell'assistente):

In questo caso, se il sistema compie un lavoro di 1000 J, deve anche assorbire un calore totale di 1000 J per soddisfare il primo principio.

mentre aggiunge al teorema di Carnot la seconda considerazione

Nel caso di tre serbatoi, il calcolo dell'efficienza diventa più complesso e dipende dal modo specifico in cui il sistema interagisce con ciascun termostato.

Sono due affermazioni essenzialmente errate. Nella prima, confonde il calore scambiato (somma dei calori assorbiti e ceduti con segno) con quello assorbito. Nella seconda, non riconosce la più ampia applicabilità del teorema di Carnot a sistemi che operano tra più di due termostati. Quindi fonde assieme—in sequenza— le due considerazioni per concludere in questo modo:

Dato che il lavoro eseguito dal sistema è di 1000 J, per il primo principio, il calore totale assorbito deve essere anch'esso di 1000 J, non 1600 J. Inoltre, il secondo principio impone che non tutto il calore assorbito possa essere convertito in lavoro. La presenza di tre sorgenti termiche a temperature diverse introduce complessità aggiuntive, ma non altera il fatto fondamentale che non tutto il calore assorbito può essere convertito in lavoro.

In conclusione, secondo i principi della termodinamica, un sistema che esegue un lavoro di 1000 J non può assorbire un calore di 1600 J in un ciclo completo, poiché ciò violerebbe il primo principio della termodinamica.

È vero che il ciclo è impossibile, ma non viola il primo principio! Sugeriamo allora all'assistente di porre attenzione al fatto che il calore non è solo assorbito ma può essere anche ceduto. ChatGPT ammette che questo è necessario e conclude correttamente che

il bilancio netto di calore assorbito e ceduto deve essere anch'esso di 1000 J:

$$Q_{\text{netto}} = W_{\text{netto}} = 1000 \text{ J.}$$

Passa poi a considerare il secondo principio e afferma:

Il calore totale assorbito dal sistema può essere maggiore di 1000 J, a condizione che una parte di questo calore venga ceduta ai termostati. Per esempio, il sistema potrebbe assorbire 1600 J da un termostato e cedere 600 J agli altri due, mantenendo un bilancio netto di 1000 J.

In realtà il calore assorbito deve essere maggiore di 1000 J, non è solamente una possibilità. Però l'errore arriva quando il ragionamento prosegue chiamando in causa il teorema di Carnot ancora maldestramente, cioè limitandone l'applicazione alla macchina bi-termica:

il teorema di Carnot stabilisce il rendimento massimo che un motore termico può raggiungere operando tra due serbatoi termici. Tuttavia, in presenza di tre serbatoi termici, come nel tuo caso, il calcolo del rendimento diventa più complesso e dipende specificamente da come il ciclo è organizzato e da come il sistema interagisce con ogni termostato. In conclusione, è teoricamente possibile per il sistema assorbire un calore totale di 1600 J in un ciclo, a condizione che ceda una parte di questo calore (in questo caso 600 J) ai termostati, mantenendo un bilancio netto di 1000 J in linea con il lavoro svolto, come richiesto dal primo principio della termodinamica. Questo scenario rispetta anche il secondo principio, poiché non implica che tutto il calore assorbito sia convertito in lavoro.

Quindi, in questa seconda versione, il ciclo proposto è considerato permesso perché rispetterebbe entrambe i principi. Insistiamo a ricordare all'assistente che il teorema di Carnot si utilizza anche con molti serbatoi. ChatGPT riconosce questa possibilità dichiarando esplicitamente che

Il rendimento di ciascun segmento del ciclo dipenderà dalle temperature dei termostati coinvolti in quello specifico scambio. In generale, la massima efficienza è raggiunta quando il sistema opera tra i due termostati con la maggiore differenza di temperatura.

Di questa affermazione però non se ne fa nulla e conclude esattamente come prima. A questo punto chiediamo esplicitamente di calcolare il rendimento "reale" e di confrontarlo con quello di Carnot. ChatGPT giunge al risultato corretto e alla conclusione che

L'efficienza massima teorica (η_{\max}) secondo il teorema di Carnot per una macchina termica operante tra i termostati a 0 °C e 400 °C è del 59.4%. L'efficienza reale (η_{reale}) del ciclo descritto nel tuo esempio, dove il sistema compie un lavoro di 1000 J e assorbe 1600 J di calore, è del 62.5%.

Interessante notare che l'efficienza reale del ciclo è leggermente superiore all'efficienza massima teorica prevista dal teorema di Carnot. Questo suggerisce che, secondo i principi della termodinamica classica, il ciclo descritto non è realizzabile nella pratica, in quanto violerebbe il secondo principio della termodinamica. In un ciclo reale, l'efficienza non può superare l'efficienza massima teorica data dal teorema di Carnot.

La risposta è finalmente giusta ma, piuttosto chiaramente, senza l'intervento dell'intelligenza naturale l'assistente virtuale non ce l'avrebbe fatta. Si è passati da un "il

ciclo è impossibile perché viola il primo principio” a un “il ciclo è possibile perché è d’accordo con tutti due i principi” prima di arrivare all’interpretazione corretta, cioè che il ciclo sia impossibile perché non è d’accordo con il secondo principio. Questo è un primo ma inquietante esempio di un’interattività con elevati potenziali di approfondimento purché i risultati corretti della questione siano chiari fin dall’inizio: è ovvio che, in caso contrario, la situazione si complica parecchio, a meno di non fidarsi ciecamente del procedimento e della dialettica artificiale e di giungere comunque lontani dalla soluzione esatta del problema. In altre parole: una persona non al corrente delle giuste modalità di porsi domande e di utilizzare i principi della termodinamica non avrebbe trovato giovamento da questa interfaccia o, per essere più espliciti, sarebbe stata indotta in clamoroso errore. Prima però di generalizzare e di trarre conclusioni, consideriamo qualche altro esempio.

La legge di aumento dell’entropia totale in un processo irreversibile, cioè non quasi-statico e/o in presenza di attriti, è ben recitata dai libri di testo e, di conseguenza, dagli studenti alle prese con il secondo principio della termodinamica. Quando si tratta di passare ai fatti, ovvero alle sue applicazioni, la situazione si complica più frequentemente di quanto non si possa immaginare. Ci sono vari aspetti che rendono la questione spesso delicata concettualmente e quindi anche negli aspetti pratici. Il fatto che l’entropia sia una funzione di stato, più nello specifico, invece di condurre a evidenti vantaggi nella risoluzione di problemi anche complessi, è trascurato, nei migliori dei casi, se non addirittura travisato nei suoi effetti. Può dunque essere interessante valutare le modalità interpretative di un assistente virtuale alle prese con un caso di studio tutto sommato tradizionale e non eccessivamente laborioso. Il testo qui proposto deriva da un esercizio che si trova in molte varianti su molti manuali, incluso nel volume dello Zemansky [22]. Si tratta di calcolare la variazione di entropia dell’universo termodinamico del processo nel quale un oggetto materiale sprofonda in un lago fino a fermarsi sul fondo. L’oggetto ha una massa e una capacità termica totale assegnate. Il lago ha una profondità nota e anche la sua capacità termica è un dato del problema. Si sa inoltre che, inizialmente, l’oggetto (un sasso, per esempio) è trattenuto fermo a pelo dell’acqua e da quella posizione lasciato cadere. La temperatura iniziale dell’acqua del lago e quella dell’oggetto sono dati del problema noti. Alla fine del processo, il sasso si ferma sul fondo del lago e con esso si riporta nelle condizioni di equilibrio termico, ovvero alla stessa temperatura finale.

Prima di avviare la conversazione con ChatGPT anticipiamo, riassumendoli, i passaggi che si considerano appropriati e corretti per la risoluzione della questione o, quantomeno, come tipicamente si possa ragionare “in modo naturale” al cospetto di un problema di questo tipo. Indichiamo con M la massa dell’oggetto e con H la profondità del lago. Denotando con “ A ” l’acqua del lago e con “ B ” l’oggetto, scriviamo le capacità termiche di A e B come C_A e C_B e con T_A e T_B le loro temperature iniziali, prima dello sprofondamento. Aspettiamoci, piuttosto ovviamente, che prima o poi dovremo tenere conto della relativamente enorme capacità termica del lago (riferita a quella dell’oggetto) e dunque nei risultati dovremo imporre la condizione $C_A \gg C_B$. Anzitutto conviene applicare il primo principio (conservazione dell’ener-

gia) nella forma $U_i^{(A)} + U_i^{(B)} + MgH = U_f^{(A)} + U_f^{(B)}$, avendo introdotto le energie interne U iniziali e finali di A e di B e avendo preso il fondo del lago ($H = 0$) come riferimento per l'energia potenziale gravitazionale di B . La variazione di energia interna totale del sistema quindi ammonta proprio alla variazione di energia potenziale gravitazionale e deve essere anche pari al calore totale trasferito al sistema, ovvero $\Delta U = MgH = Q^{(A)} + Q^{(B)}$. I calori sono collegati alla variazione di temperatura di A e B espressa tramite la relazione che si assume essere qui applicabile, $Q = C\Delta T$ e dunque, in definitiva, si ottiene che

$$Q = Q^{(A)} + Q^{(B)} = C_A(T_F - T_A) + C_B(T_F - T_B) = MgH,$$

dalla quale si ricava l'espressione per la temperatura di equilibrio finale T_F :

$$T_F = \frac{MgH}{C_A + C_B} + \frac{C_A T_A + C_B T_B}{C_A + C_B}.$$

Queste espressioni possono ovviamente essere valutate in casi particolari. Se si parla, per esempio, del lago con capacità termica pressoché illimitata, la temperatura finale risulta pari a quella posseduta inizialmente dal lago stesso, $T_F = T_A$. In questa situazione, si osserva quindi che il lavoro svolto dalla forza peso (MgH) non è in grado di influenzare apprezzabilmente la temperatura di equilibrio finale, cosa che invece accade se le capacità termiche di A e B sono in qualche misura confrontabili. Se poi, per esempio, A e B sono inizialmente alla stessa temperatura T_i , la temperatura di equilibrio risulta pari a $T_F = T_i + MgH/(C_A + C_B)$. Restando ancora nell'ambito del bilancio energetico, si ottengono subito le espressioni per i calori scambiati da A e B :

$$Q^{(A)} = \frac{C_A C_B}{C_A + C_B} (T_B - T_A) + \frac{C_A}{C_A + C_B} MgH,$$

$$Q^{(B)} = \frac{C_A C_B}{C_A + C_B} (T_A - T_B) + \frac{C_B}{C_A + C_B} MgH.$$

Si nota, in particolare, la distribuzione dell'energia potenziale gravitazionale iniziale, MgH , a favore dell'oggetto con maggiore capacità termica: nell'ipotesi che A abbia una capacità termica molto maggiore di quella di B si ottiene infatti che

$$Q^{(A)} \approx C_B(T_B - T_A) + MgH \quad \text{e} \quad Q^{(B)} \approx C_B(T_A - T_B).$$

L'ultimo passo è relativo all'applicazione del secondo principio nella forma del calcolo della variazione di entropia dell'universo termodinamico che, nelle presenti ipotesi di lavoro, è costituito esclusivamente dagli oggetti A e B , ovvero non si considerano l'aria e altri elementi esterni (il che complicherebbe inutilmente la situazione). Considerando il fatto che l'entropia è funzione di stato e che essa è definita a partire dal calore

elementare scambiato reversibilmente alla temperatura T , nel caso in esame è sufficiente integrare tra gli stati iniziale e finale il differenziale esatto $dS = \delta Q/T = C dT/T$ il che conduce al risultato

$$\Delta S_{\text{univ}} = \Delta S_A + \Delta S_B = C_A \log(T_F/T_A) + C_B \log(T_F/T_B).$$

A partire da questa espressione, di validità essenzialmente generale basandosi sulle assunzioni di questo problema, si possono considerare dei casi particolari, come già fatto nel calcolo della temperatura di equilibrio finale e dei calori scambiati. Se, per esempio, all'inizio c'è già equilibrio termico, $T_i = T_A = T_B$, si ottiene immediatamente che entrambe gli oggetti A e B danno contributi entropici non nulli ed entrambi positivi causati dalla dissipazione dell'energia potenziale gravitazionale:

$$\Delta S_{\text{univ}} = C_A \log \left[\frac{1}{T_i} \frac{MgH}{C_A + C_B} + 1 \right] + C_B \log \left[\frac{1}{T_i} \frac{MgH}{C_A + C_B} + 1 \right].$$

Si osserva in questo risultato che, se la capacità termica di A è molto maggiore di quella di B (il caso del lago), allora il termine entropico dell'oggetto B tende ad annullarsi mentre rimane non nullo quello dell'oggetto A (è sufficiente utilizzare l'approssimazione $\log(x+1) \cong x$ per $x \ll 1$). Esplicitamente quindi la variazione dell'entropia totale si semplifica al valore

$$\Delta S_{\text{univ}} = \Delta S_A = \frac{MgH}{T_i}.$$

L'oggetto a bassa capacità termica (il sasso nel lago) non aumenta apprezzabilmente la sua entropia perché il suo stato termodinamico non varia complessivamente nel processo. L'entropia riceve contributo solo dal trasferimento irreversibile di energia gravitazionale in forme temporanee meccanicamente complesse (turbolenze, vortici, onde, ecc.) all'acqua del lago. È interessante notare che, in caso di assenza iniziale di equilibrio termico tra A e B (oggetto e lago a diverse temperature iniziali) l'entropia riceve un contributo di irreversibilità anche termico che è subito calcolato sempre a partire dalla medesima espressione generale sopra ottenuta. Il risultato, sempre nell'ipotesi $C_A \gg C_B$, è

$$\Delta S_A = \frac{MgH}{T_A}, \quad \Delta S_B = \frac{C_B}{C_A} \frac{MgH}{T_A} + C_B \log \left(\frac{T_A}{T_B} \right);$$

in questa espressione va fatto notare che T_A è anche la temperatura finale di equilibrio del sistema composto dal lago e dal sasso mentre T_B è la temperatura iniziale del sasso. La variazione totale di entropia è quindi

$$\Delta S_{\text{univ}} = \frac{MgH}{T_F} + C_B \log \left(\frac{T_F}{T_B} \right)$$

in cui $T_F = T_A$. Pur nella sua semplicità, questo risultato è —anche didatticamente— notevole perché include allo stesso tempo i due enunciati di irreversibilità termodinamica che realizzano il secondo principio: c'è infatti la parte di irreversibilità meccanica

(qui sempre positiva) e quella di non-staticità per salto termico (che qui può essere sia negativa che positiva). Passando l'analisi della situazione a ChatGPT i risultati non sono incoraggianti. La nostra domanda iniziale:

un oggetto con capacità termica totale costante C_B e massa M è inizialmente alla temperatura T_B . Viene lasciato affondare, partendo da fermo, in un lago con capacità termica costante C_A che si trova alla temperatura T_A e di profondità H . L'oggetto giunto sul fondo del lago si ferma e tutto torna alla fine all'equilibrio termico. Calcola la variazione di entropia totale di questo processo quando $C_A \gg C_B$.

Il primo tentativo è basato sulla versione pubblica dell'interfaccia, la 3.5, al momento di scrivere. La risposta è un addensato di errori con davvero poche possibilità di perdono. Riassumendo il flusso principale della conversazione (i punti esclamativi e le parti non in corsivo sono dell'autore):

- (a) *per il lago il processo viene visto come adiabatico perché la sua variazione di temperatura è molto piccola (!);*
- (b) *dunque, non essendoci scambi di calore significativi, la variazione di entropia del lago è nulla (!);*
- (c) *per l'oggetto si calcola il calore scambiato (con chi?) come $Q = MC\Delta T$ (in realtà, nel testo si parla di capacità termica, non di calore specifico) che è subito calcolato perché si sa la variazione di temperatura dell'oggetto;*
- (d) *la variazione di entropia dell'oggetto si ottiene a partire dalla sua definizione, $\Delta S = Q/T$ (!);*
- (e) *quella dell'universo è data quindi da quella dell'oggetto appena calcolata.*

Prima di protestare, chiediamo all'assistente cosa succede, in base al suo procedimento, se la temperatura del sasso e del lago sono le stesse fin dall'inizio. La risposta, se non altro, è coerente con il risultato, ovvero:

La variazione di entropia totale del sistema sarà determinata principalmente dalla variazione di entropia del lago (ΔS_A). Se il lago è considerato un sistema termalmente molto grande e stabile, la sua variazione di entropia potrebbe essere trascurabile o molto piccola, a meno che non ci siano altri processi significativi che coinvolgano il lago. In generale, se le temperature iniziali di due sistemi sono uguali, non ci sarà scambio di calore tra di loro, e la variazione di entropia complessiva dipenderà da altri fattori o processi presenti nel sistema.

Chiediamo di tenere conto fin dall'inizio di questi altri fattori, chiamandoli sorgenti di irreversibilità meccanica. ChatGPT risponde

Hai ragione, chiedo scusa per la mia mancanza di chiarezza. Nel problema in esame, dovremmo effettivamente considerare l'irreversibilità meccanica fin dall'inizio, in quanto l'oggetto si ferma quando raggiunge il fondo del lago e

il moto è stato arrestato. In un sistema reale, il moto dell'oggetto sarebbe stato dissipato sotto forma di lavoro contro le forze di resistenza, come il drag dell'acqua.

Pertanto, anche se l'entropia totale del sistema rimane costante ($\Delta S_{\text{totale}} = 0$), il processo è irreversibile dal punto di vista meccanico. Questo significa che l'energia meccanica è stata dissipata sotto forma di lavoro contro le forze di resistenza durante il processo, rendendo impossibile riportare l'oggetto esattamente nello stesso stato iniziale senza fornire energia esterna aggiuntiva.

L'asserire che: “anche se l'entropia non cambia” il processo è irreversibile, è piuttosto bizzarro. Però proviamo con un altro suggerimento, ovvero di ricordare che l'entropia cambia anche a seguito di fenomeni dissipativi di ogni genere. La risposta è definitiva:

Hai assolutamente ragione. La variazione di entropia misura anche il grado di irreversibilità meccanica in un processo termodinamico. In un processo irreversibile, la variazione di entropia del sistema aumenta, mentre in un processo reversibile, la variazione di entropia è zero o rimane costante. Quindi, se la variazione di entropia è diversa da zero, indica un certo grado di irreversibilità meccanica nel processo.

Nel problema iniziale, se la variazione di entropia totale è zero ($\Delta S_{\text{totale}} = 0$), ciò indica che il processo non comporta alcuna variazione netta di entropia e che il sistema è tornato a un equilibrio termodinamico, ma questo non implica necessariamente che il processo sia meccanicamente reversibile. In realtà, come hai correttamente sottolineato, il processo può ancora essere irreversibile meccanicamente, il che significa che l'energia meccanica è stata dissipata sotto forma di lavoro contro le forze di resistenza, anche se l'entropia totale rimane costante.

Insomma, secondo ChatGPT questo è un processo irreversibile (e ci mancherebbe) ma l'entropia può benissimo non cambiare. Il che lascia un po' a desiderare, come risposta “intelligente”: il contributo di irreversibilità meccanica misurato proprio dal valore di MgH riferito alla temperatura di equilibrio termico è completamente scomparso. Non va meglio con la versione “pro” di ChatGPT (la 4.0). La domanda viene formulata esattamente nello stesso modo di prima e la risposta è in linea con la precedente ovvero, anche se le idee di base sull'entropia sono formulate abbastanza correttamente, il bilancio energetico è errato e questo si ribalta inevitabilmente sul calcolo dell'ammontare di irreversibilità. ChatGPT dichiara essenzialmente che

- (a) *la variazione totale di entropia è divisa in quella dell'oggetto e in quella del lago;*
- (b) *la variazione di entropia di un oggetto che cambia temperatura da T_i a T_F può essere calcolata usando la formula $C \log(T_F/T_i)$;*
- (c) *nel caso del lago, siccome non cambia la sua temperatura ma scambia una quantità di calore Q (usare questa terminologia per una modalità di trasferimento energetico è didatticamente deprecabile) con l'oggetto, la sua variazione di entropia sarà $C(T_i - T_F)/T_i$.*

Notiamo che la risposta (b) è quantomeno traballante —se un gas ideale si trasforma reversibilmente in modo adiabatico la sua temperatura cambia ma la sua entropia no. Però, a questo punto, si chiede a quanto ammonterebbe la variazione totale di entropia nel caso in cui il sasso e il lago siano già all'equilibrio termico fin dall'inizio. La risposta è lapidaria:

il valore di variazione di entropia è zero perché non avviene scambio netto di calore tra l'oggetto e il lago.

Chiediamo: “quindi il processo è reversibile?” e leggiamo che

La variazione di entropia è zero significa semplicemente che non c'è stato scambio di calore e non perché il processo sia reversibile . . . mentre l'assenza di variazione di entropia in questo caso specifico suggerisce un processo teoricamente reversibile, la realtà fisica di tale processo, considerando tutti i fattori, è tipicamente irreversibile.

Chiediamo se è possibile includere il contributo del fattore gravitazionale (il lavoro fatto dalla forza peso) e l'assistente ammette che questo è un aspetto rilevante dell'analisi (anche se lo definisce un qualcosa che aggiunge ulteriore complessità al problema). Riconosce che l'energia dissipata è pari a MgH . Poi però aggiunge che

Una parte di questa energia potrebbe essere convertita in energia termica, che aumenta la temperatura dell'oggetto e, possibilmente, quella del lago. Tuttavia, il dettaglio esatto di come questa energia si divide tra l'oggetto e l'acqua del lago (e quanta di essa si disperde in altre forme come l'energia cinetica dell'acqua) dipende dalla dinamica specifica dell'affondamento, che non è stata dettagliata nel problema originale.

Con un ultimo tentativo, richiamiamo ancora l'attenzione sul fatto che l'entropia è una funzione di stato e quindi non sono necessari questi dettagli della dinamica specifica dell'affondamento per il calcolo della variazione dell'entropia stessa. ChatGPT è d'accordo ma conclude ancora una volta che

Se consideriamo il sistema nel suo insieme, la variazione di entropia totale dovuta al trasferimento di calore dal rilascio di energia potenziale gravitazionale può essere approssimata considerando il calore totale assorbito e la temperatura a cui il calore viene assorbito. Tuttavia, senza dettagli specifici su come questo calore viene distribuito (quanto viene assorbito dall'oggetto e quanto dall'acqua del lago), non possiamo calcolare esattamente la variazione di entropia.

Il punto chiave è che, anche includendo l'energia gravitazionale, l'analisi dell'entropia rimane centrata sulle condizioni finali e iniziali del sistema. La variazione di entropia totale rifletterà tutti i cambiamenti energetici, compresa la conversione dell'energia potenziale gravitazionale in calore, purché si possano definire chiaramente queste condizioni.

In pratica, niente da fare: l'esercizio non viene risolto o, peggio ancora, traspare neanche troppo nascostamente il totale distacco, secondo l'assistente virtuale, tra principio di aumento dell'entropia e irreversibilità del processo: si può ben dire che, nell'opinione di ChatGPT, il secondo principio della termodinamica non funziona in questo problema. Anche il tentativo di "guidare" il l'assistente verso la soluzione corretta passando attraverso il calcolo della temperatura di equilibrio prima e i calori scambiati poi, come è stato accennato poco sopra, non conduce a risultati migliori: anzitutto ChatGPT non riesce a mettere correttamente a bilancio l'energia del processo, pur partendo dall'ipotesi della sua conservazione. Quindi fallisce nell'intento di fornire la temperatura di equilibrio. I calori risultano di conseguenza errati e, in cascata, anche l'entropia finale non viene espressa nella forma prevista. Inoltre, i tentativi di portare all'attenzione dell'assistente dei casi particolari utili a chiarire la situazione (utilizzare per esempio uno stato iniziale già di equilibrio termico) non hanno funzionato. Le tecniche cosiddette di "*prompt injection*", ovvero di inserimento di affermazioni *ad hoc* che possano aprire o innescare nuove vie di risoluzione non sono risultate efficaci per questo problema. Si tratta di un caso di studio che, anche se non completamente banale, dovrebbe essere accessibile a livello concettuale e operativo a chiunque si voglia appropriare di una competenza minima e irrinunciabile nel campo dei fondamenti della termodinamica classica, come quella affrontata in un buon percorso liceale o nei corsi di fisica generale a vari indirizzi universitari.

3. I concetti virtuali

Si potrebbe continuare a produrre esempi delle difficoltà nelle quali incappa ChatGPT (3.5 o 4.0 non importa, ma anche Gemini/Bard e Copilot non fanno meglio) nel tentativo di risolvere varie tipologie di esercizi e di problemi nei quali è necessario applicare —con molta cura, non in modo meccanico e automatico— i principi e le leggi della termodinamica classica. Allo scopo di indagare anche altri aspetti della questione qui in esame, è interessante proporre delle domande più "dirette", concettuali, meno applicative, sempre restando nel campo della termodinamica. Si scopre allora che ChatGPT riesce con maggiore frequenza a reagire correttamente producendo risposte che effettivamente sono spendibili come aiuto allo studio, al ripasso, a considerare eventuali approfondimenti. Con moltissime cautele, però: le insidie della termodinamica si fanno sentire quando, per esempio, poche o addirittura singole parole possono fare la differenza tra la risposta giusta e le più assurde e strampalate interpretazioni, oppure a generare quelle che in gergo sono chiamate "allucinazioni" dell'assistente [23].

Chiediamo, per esempio, se è vero che il primo principio sia applicabile solo a processi reversibili. Qui, ChatGPT va a colpo sicuro e rassicura sul fatto che la legge di conservazione dell'energia è universale e si applica anche a processi non reversibili. Testualmente:

No, il primo principio della termodinamica è applicabile sia ai processi reversibili che a quelli irreversibili. Il primo principio, noto anche come legge della conservazione dell'energia, afferma che l'energia di un sistema isolato è costante. In altre parole, l'energia non può essere creata né distrutta, ma può solo trasformarsi da una forma all'altra o trasferirsi da un sistema all'altro.

Non si fa nemmeno ingannare con la domanda abbastanza popolare tra i libri di testo relativa all'opportunità di tenere acceso il ventilatore in una stanza chiusa durante una calda giornata estiva, allo scopo di rinfrescare l'ambiente. La risposta non è ovvia e immediata: certamente avere un motore elettrico acceso nella stanza per far funzionare il ventilatore non è una buona idea se l'intenzione è quella di diminuire la temperatura ambientale, visto che il motore è una macchina che si surriscalda in qualche misura. C'è però da considerare il fatto che l'aria smossa dal ventilatore è in grado di generare una sensazione di refrigerio dovuta, per esempio, all'aumentata evaporazione del sudore epidermico (*evaporative cooling* [24]), fenomeno questo facilitato dal moto dell'aria che incrementa il trasporto termico noto come "convezione forzata". Su questi aspetti ChatGPT rivela una buona preparazione ovvero un addestramento di qualità. Testualmente:

Sì, tenere in funzione un ventilatore in una stanza chiusa durante una calda giornata d'estate può essere utile per rinfrescarsi (...). Se la pelle è umida (per esempio, a causa del sudore), l'aria mossa dal ventilatore può accelerare l'evaporazione del sudore, sottraendo calore dal corpo e facendo sentire più freschi. Tuttavia, è importante notare che un ventilatore non riduce effettivamente la temperatura dell'aria; piuttosto, il movimento dell'aria può migliorare la percezione del comfort.

Alla domanda invece se sia corretto affermare che il secondo principio della termodinamica escluda la possibilità che in una trasformazione sia possibile convertire interamente un certo ammontare di calore in lavoro, la risposta è poco convincente o, quanto meno, ambigua. Nella formulazione appena riportata, è chiaro che esistono trasformazioni perfettamente in grado di realizzare il processo descritto: è sufficiente prendere in considerazione un gas ideale che, reversibilmente, si espande a contatto con un serbatoio termico a temperatura fissata. Si tratta di una trasformazione isoterma che, notoriamente, converte tutto il calore assorbito dal termostato in lavoro di volume idrostatico che l'espansione del gas consente. ChatGPT invece esordisce con un secco "sì, questo è quanto prevede il secondo principio della termodinamica". Fa sfoggio delle sue conoscenze aggiungendo che questa impossibilità è osservata in processi ciclici nei quali non ci sia traccia di altri cambiamenti nel sistema. Il che è corretto, ma nella domanda posta non si parlava di questo tipo particolare di processi ma di "una trasformazione", senza ulteriori specifiche. Queste le ha aggiunte il chatbot distorcendo in questo modo la visione più ampia del problema. Quando viene fatto notare questo fatto, la risposta che si riceve è che nell'esempio di una trasformazione isoterma singola si considera solamente una parte di un ciclo e non di un

processo complessivo che riporta il sistema allo stato iniziale. Cosa che non era stata chiesta. Se si insiste con questo commento, ChatGPT ringrazia della puntualizzazione ma non sposta di un millimetro la sua risposta iniziale, cioè che il secondo principio non permette “trasformazioni” che riescono a fare invece benissimo quanto detto. In questo caso la difficoltà è forse più di origine semantica che altro, ma questo preoccupa ancora di più, visto che l’assistente dovrebbe proprio eccellere in situazioni nelle quali sono le costruzioni di linguaggio più che di comprensione tecnica a prevalere.

Un ultimo esempio di funzionamento contraddittorio di ChatGPT prende spunto dalla domanda nella quale si chiede se, partendo da un sistema fisico con stati iniziale e finale perfettamente determinati, anche la variazione di entropia sia fissata indipendentemente dalla presenza di attriti o salti termici finiti. In prima battuta, ancora una volta, ChatGPT prende tempo con una serie verbosa di considerazioni che però sono accomunate da una bizzarra idea di entropia: nel caso di processi reversibili, la variazione di S è certamente indipendente dal processo seguito per collegare gli stati iniziale e finale; però, in un processo “reale” o irreversibile, la variazione di entropia sarebbe maggiore di quella che collega gli stessi stati in una trasformazione reversibile. Può essere che questo schema descrittivo in realtà voglia spiegare che la variazione di entropia dell’universo termodinamico (e non del sistema!) di un processo reversibile è sempre nulla, mentre quella per il processo irreversibile che collega gli stessi stati è sicuramente positiva. Ma questa non è la domanda posta. Alla richiesta di specificare se l’entropia è una funzione di stato (chiave di lettura essenziale in questo contesto), ChatGPT risponde correttamente, aggiungendo una lunga e non richiesta spiegazione, secondo il suo stile che, a lungo andare, può essere piuttosto irritante. A questo punto si fa notare che il fatto di essere funzione di stato automaticamente conduce alla risposta che contraddice quella fornita poche righe di conversazione sopra. Allora ChatGPT comincia a chiarire, apparentemente più a sé stessa che a chi la interroga, che bisogna distinguere tra sistema e ambiente e tra variazione e produzione di entropia. Al netto di tutto ciò, quando si insiste per un’ultima volta su quale sia in definitiva la risposta alla domanda originale, ChatGPT si arrende e ammette che “sì, la variazione di entropia dipende solamente dagli stati di partenza e arrivo indipendentemente dalla presenza di attriti o salti termici finiti”. Meglio tardi che mai.

4. Qualche riflessione e conclusioni

L’esplorazione della termodinamica — e di altri campi delle scienze fisiche, in contesti didattici o più generalmente di ricerca — avvalendosi di questi LLM può ovviamente continuare senza virtualmente nessun limite. Sistematicamente, però, si evidenziano grossi limiti legati sia a questioni di natura tipicamente semantica che ad aspetti di incorretta o incompleta concatenazione di relazioni causa-effetto o, meglio, di teoria e pratica relativamente ai contenuti della disciplina. Per quanto riguarda questo lavoro di osservazione e sperimentazione preliminare sul vastissimo tema della didattica della termodinamica classica, si sono evidenziati degli impedimenti anche gravi, una

bassa affidabilità e prestazioni molto variabili nelle abilità risolutive di problemi concreti, dove contano ovviamente i principi più generali ma anche, e soprattutto, la loro applicazione, la loro contestualizzazione, la manipolazione simbolica di espressioni di complessità non necessariamente elevatissima e la traduzione di relazioni astratte in valori numerici di determinate grandezze fisiche. Tutto ciò è tanto più critico quanto minore è l'intervento "umano", basato sulle azioni di "smart prompt", cioè di guida passo-passo per raggiungere il risultato corretto, che dunque già deve essere noto a priori assieme alla strategia più efficace di risoluzione. In questo periodo del suo sviluppo, risulta dunque essenziale e strategico un continuo e articolato lavoro di miglioramento, monitoraggio, sperimentazione controllata su campioni di popolazione studentesca selezionata e limitata. Prima di passare a una "sperimentazione di massa" è di drammatica rilevanza pedagogica, didattica e motivazionale individuare al meglio i limiti di operatività degli LLM. Si può certamente discutere, purché lo si faccia a fronte di dati statisticamente robusti e quindi non in tempi troppo ravvicinati e affrettati, sull'opportunità generale di affiancare, se non addirittura sostituire le azioni concettuali che l'intelligenza naturale mette in atto durante il suo sviluppo scolastico. Il problema è certamente complesso e le risposte non sono dietro l'angolo perché mancano ancora le giuste domande da porre. Certo è che non è prudente lasciare che questi assistenti entrino, incontrollati, nel mondo della scuola (di qualsiasi ordine si tratti). Si potrebbero riassumere i possibili ruoli immaginabili in un contesto didattico e dell'istruzione in generale che un assistente virtuale dovrebbe assicurare in una certa misura per essere "assunto" in questo impiego: anzitutto una funzione di vero e proprio istruttore, o insegnante permanente e artificiale. Certamente non allo stato attuale, viste le prestazioni che questi linguaggi stanno fornendo. In particolare, come già accennato, è molto pericoloso il voler considerare "dialettico" o "maieutico" il colloquio che si può instaurare con un tutore di scarse conoscenze "didatticamente utili", se non addirittura contenutisticamente errate. Non è molto efficace prendere lezioni di guida da qualcuno che non è mai salito su un'automobile. Non si capisce cosa si potrebbe apprendere di utile (oltre alla frustrazione del tempo perduto) dialogando di termodinamica con un assistente pronto a sbagliare completamente i suoi ragionamenti senza che lo studente sappia di questi errori né sia a conoscenza del risultato "corretto" da ottenere. Forse potrebbe essere più interessante immaginare che ChatGPT si comporti come un compagno o una compagna di studio: in questi casi è ammessa la possibilità di non sapere la risposta giusta e di avviare una discussione anche ispiratrice di nuove idee quando si deve affrontare un problema da zero. Si deve però tenere presente la natura fortemente e inevitabilmente "non naturale" delle affermazioni dell'assistente: con ciò si intende sottolineare che le modalità di interrogazione e di revisione dei propri modi di ragionare, dubitare, impostare, risolvere situazioni e problemi alla richiesta di risposte, quando costruite su schemi e meccanismi naturali, anche se non possono attingere a una sterminata quantità di dati sicuramente rilevanti come possono fare i motori GPT, sono tuttavia capaci di correlare e di "vedere oltre" alla conoscenza di "cose" per arrivare alla conoscenza di "come". Non sappiamo bene (spesso neanche un po') come definire e dunque comprendere come si comprende (o non si comprende). Però ci è abbastanza chiaro che

la quantità di informazioni non sono tutto. Il nostro assistente virtuale, per ora almeno, pur stracolmo di bit di “cosa”, fatica a dialogare con le nostre visioni molto più imprevedibili e creative, critiche, sorprendenti proprio per queste caratteristiche. Probabilmente vale anche la pena di riflettere sulle possibilità che gli LLM hanno di fornirci materiale utile alla valutazione degli studenti “veri”, costruendo questionari, testi di quiz, prove scritte, assieme alle procedure automatizzate ed estremamente rapide per l’attribuzione di punteggi e per l’analisi dei risultati. Questo apre di fatto anche il campo di sperimentazione per l’aggiornamento e la formazione iniziale e in itinere degli e delle docenti di ogni ordine di scuola, non solo nell’ambito delle scienze fisiche.

Questo lavoro, necessariamente incompleto nel senso della mancanza di conclusioni definitive, non intende in alcun modo sminuire l’enorme impegno dei gruppi di ricerca che negli anni hanno condotto alla costruzione e alla condivisione con il grande pubblico di questi strumenti di eccezionale potenzialità culturale e di insostituibile supporto alla complessità della nostra società. Si vuole però richiamare l’attenzione sulla drammatica importanza di non lasciar andare in ordine sparso le sperimentazioni, le valutazioni, le verifiche, gli studi a moltissimi livelli che sono necessari per mettere in ordine fin da ora, prima che sia troppo tardi, le implicazioni e le applicazioni dell’IA nel mondo dell’istruzione e della ricerca in didattica. L’ondata è arrivata possente, non ci si deve far travolgere. Bisogna cavalcarla, anche se non sarà facile. È fondamentale non far passare l’idea e l’atteggiamento che “visto che c’è, sfruttiamola”, un po’ come si è fatto con le calcolatrici prima, i computer a seguire, e altro ancora più sofisticato in sequenza. Prima di sfruttare una tecnologia —e le IA sono tali, vanno viste e riconosciute come strumenti, non come menti pensanti!— è necessario conoscerne le caratteristiche, le modalità di utilizzo e i limiti, a maggior ragione quando si intende utilizzarle in campo educativo e didattico. Come già fatto notare in vari passi di questo lavoro, allo stato attuale dell’arte questi parametri di misura delle IA non sono noti a sufficienza o, addirittura, sono del tutto indeterminati e non valutabili. Inoltre, lo si rimarca ma forse mai abbastanza, il continuo delegare e trasferire mestieri di manovalanza a “basso livello” a una servitù virtuale è una strategia spesso perdente sul fronte della crescita delle capacità e dell’impegno cognitivo. I docenti di ogni ambito educativo non dovrebbero mai sottovalutare questo rischio.

Bibliografia

- [1] POLVERINI G. and GREGORCIC B., “How understanding large language models can inform the use of ChatGPT in physics education”, *Eur. J. Phys.*, **45** (2024) 025701.
- [2] YEADON W., INYANG O.-O., MIZOURI A., PEACH A. and TESTROW C. P., “The death of the short-form physics essay in the coming AI revolution”, *Phys. Educ.*, **58** (2023) 035027.
- [3] DAHLKEMPER M. N., LAHME S. Z. and KLEIN P., “How do physics students evaluate artificial intelligence responses on comprehension questions? a study on the perceived scientific accuracy and linguistic quality of ChatGPT”, *Phys. Rev. Phys. Educ. Res.*, **19** (2023) 010142.
- [4] GREGORCIC B. and PENDRILL A.-M., “ChatGPT and the frustrated Socrates”, *Phys. Educ.*, **58** (2023) 035021.
- [5] OSS S., “Artificial intelligence at school: please handle with care”, *Phys. Educ.*, **58** (2023) 046501.

- [6] KORTEMEYER G., “Could an artificial-intelligence agent pass an introductory physics course?”, *Phys. Rev. Phys. Educ. Res.*, **19** (2023) 010132.
- [7] YEADON W., INYANG O.-O., MIZOURI A., PEACH A. and TESTROW C. P., “The death of the short-form physics essay in the coming AI revolution”, *Phys. Educ.*, **58** (2023) 035027.
- [8] REZENDE M. F. JUNIOR and LÓPEZ-SIMÓ V., “What are the perceptions of physics teachers in Brazil about ChatGPT in school activities?”, *J. Phys.: Conf. Ser.*, **2693** (2024) 012011999.
- [9] BITZENBAUER P., “ChatGPT in physics education: a pilot study on easy-to-implement activities”, *Contemp. Educ. Technol.*, **15** (2023) ep430.
- [10] WEST C. G., “AI and the FCI: Can ChatGPT project an understanding of introductory physics?”, arXiv:2303.01067.
- [11] DOS SANTOS R. P., “Enhancing physics learning with ChatGPT, Bing chat, and Bard as agents-to-think-with: a comparative case study”, arXiv:2306.00724.
- [12] WEST C. G., “Advances in apparent conceptual physics reasoning in GPT-4”, arXiv:2303.17012.
- [13] YEADON W. and HARDY T., “The impact of AI in physics education: a comprehensive review from GCSE to university levels”, arXiv:2309.05163.
- [14] YEADON W. and HALLIDAY D. P., “Exploring Durham university physics exams with large language models”, arXiv:2306.15609.
- [15] Bard/Gemini Google IA, <https://gemini.google.com/>
- [16] Copilot Microsoft IA, <https://copilot.microsoft.com/>
- [17] Piattaforme Open AI, <https://platform.openai.com/docs/models/overview>
- [18] Applicazioni AI secondo Forbes, <https://www.forbes.com/sites/bernardmarr/2023/05/10/15-amazing-real-world-applications-of-ai-everyone-should-know-about/>.
- [19] LEVRINI O., FANTINI P., PECORI B. and TASQUIER G., “Forms of productive complexity as criteria for educational reconstruction: the design of a teaching proposal on thermodynamics”, *Social Behav. Sci.*, **116** (2014) 1483.
- [20] ERCOLI A. and LUBICZ V., “Termodinamica e teoria microscopica: una proposta didattica per la scuola superiore”, *G. Fis.*, **64** (2023) 223.
- [21] TALANQUER V., “Interview with the chatbot: how does it reason?”, *J. Chem. Educ.*, **100** (2023) 2821.
- [22] DITTMAN R. H. and ZEMANSKY M. W., *Heat and Thermodynamics*, VII ed. (McGraw Hill, New York, NY) 1997.
- [23] BANG Y., CAHYAWIJAYA S., LEE N., DAI W., SU D., WILIE B., LOVENIA H., JI Z., YU T., CHUNG W., DO Q. V., XU Y. and FUNG P., “A multitask, multilingual, multimodal evaluation of ChatGPT on reasoning, hallucination, and interactivity”, arXiv:2302.04023.
- [24] HUGHES S. and SCHOUTEN P., “The physics of outdoor air conditioning and evaporative cooling”, *Phys. Educ.*, **58** (2023) 035031.