

Usi intelligenti dell'Intelligenza Artificiale. Il *man-with-the-machine learning*.

Roberto Trinchero, roberto.trinchero@unito.it

Dipartimento di Filosofia e Scienze dell'Educazione, Università degli studi di Torino

Riassunto

L'uso dei sistemi di intelligenza artificiale in grado di dialogare in linguaggio naturale con gli umani e di rispondere a consegne precise sfruttando una base di conoscenza alimentata dalla rete (chatbot), si sta diffondendo anche in ambito didattico, facendo intravedere possibilità e rischi. L'articolo intende presentare alcune riflessioni legate all'utilizzo delle macchine per insegnare e alla loro evoluzione in sistemi intelligenti di tutoraggio, delineando le possibilità e i rischi di un'interazione di apprendimento/insegnamento tra l'umano e la macchina guidata dalla macchina e di una guidata dall'umano. Le riflessioni si focalizzano sui possibili principi e modalità per instaurare relazioni di apprendimento/insegnamento efficaci, in grado di valorizzare gli aspetti peculiari dell'intelligenza umana e di potenziarli con l'ausilio dell'"intelligenza" della macchina. La prospettiva del *man-with-the-machine learning* definisce e promuove modalità di interazione sinergica tra i due tipi di intelligenza – profondamente e strutturalmente diversi –, mostrando come possano agire insieme e completarsi vicendevolmente.

Abstract

The use of artificial intelligence systems capable of conversing in natural language with humans and responding to precise deliveries by leveraging a network-powered knowledge base (chatbots) is also spreading to the educational field, hinting at possibilities and risks. This paper aims to present some reflections related to the use of machines for teaching and their evolution into intelligent tutoring systems, outlining the possibilities and risks of human-human and human-guided machine learning/teaching interaction. The reflections focus on possible principles and ways to establish effective learning/teaching relationships that can enhance the unique aspects of human intelligence and empower them with the help of machine "intelligence". The *man-with-the-machine learning* perspective defines and promotes ways of synergistic interaction between the two types of intelligence - deeply and structurally different - showing how they can act together and complement each other.

Parole chiave: Intelligenza artificiale, Tecnologie per l'apprendimento, Sistemi intelligenti di tutoring, Didattica efficace, Sinergia umano-macchina.

Keyword: Artificial intelligence, Technologies for learning, Intelligent tutoring systems, Effective teaching, Human-machine synergy.

Testo:

“Studiare ogni aspetto dell'apprendimento o qualsiasi altra caratteristica dell'intelligenza in modo da descriverli così precisamente da poter costruire una macchina che li *simuli*”. Questo era l'obiettivo della proposta di John McCarthy, Marvin Minsky, Nathan Rochester e Claude Shannon del 31 agosto 1955 di istituire, per l'estate del 1956, un seminario estivo di 2 mesi presso il Dartmouth College di Hanover nel New Hampshire durante il quale gettare le basi programmatiche di una nuova disciplina volta a studiare la possibilità di creare forme di “intelligenza artificiale” (McCarthy et al. 2006). Obiettivi del seminario erano quelli di studiare in maniera approfondita le modalità con cui la mente umana riconosce termini linguistici, formula concetti, produce implicazioni tra di essi e soluzioni a problemi, riducendo lo spazio delle alternative possibili attraverso il ragionamento. Questi saperi avrebbero dovuto generare idee per sfruttare al meglio le capacità dei computer dell'epoca (e futuri) nel formulare astrazioni a partire dai dati disponibili, avviare percorsi di auto-miglioramento, sviluppare forme di pensiero creativo introducendo un grado di casualità controllato nel “pensiero” algoritmico-esecutivo che li guidava.

Gli obiettivi erano giustamente ambiziosi e hanno alimentato più filoni di ricerca che hanno portato alle evoluzioni tecnologiche che conosciamo. Quello che è importante focalizzare è però l'intento di costruire una macchina in grado di *simulare* l'intelligenza umana su ambiti e problemi ben definiti, non di costruire qualcosa in grado di *sostituirsi* ad essa. Era molto chiaro agli autori, così come a coloro che li hanno seguiti, che umani e macchine elaborano informazione in modo profondamente differente. L'elaborazione umana avviene nel mondo degli stati mentali intrapersonali (il mondo 2 di Popper-Eccles, 1981), l'elaborazione della macchina avviene nel mondo delle costruzioni simboliche (il mondo 3 di Popper-Eccles). L'umano processa informazioni attraverso funzioni cognitive che conducono alla costruzione di rappresentazioni mentali dotate di senso e significato per l'umano che li elabora, dove per *significato* si intende la loro corrispondenza con entità del mondo reale e per *sensò* le modalità con cui queste entità sono percepite e interpretate dagli individui (Anderson, 2009).

La macchina elabora simboli che per essa non hanno né senso né significato. Senso e significato all'informazione ricevuta o prodotta dalla macchina lo può assegnare solo l'umano. La forza della macchina sta nella possibilità di elaborare moli di informazione simbolica che l'umano non sarebbe in grado di elaborare in tempi ragionevoli; la forza dell'umano sta nella possibilità di assegnare senso e significato a queste informazioni. Vista questa complementarità, la macchina ha bisogno dell'umano come l'umano ha bisogno della macchina. Un uso "intelligente" dell'intelligenza artificiale parte quindi dalla risposta agli interrogativi "Come possiamo costruire sinergie ottimali tra umano e macchina 'intelligente'?", "Come possiamo costituire sistemi collaborativi 'umano-macchina' in grado di promuovere sia apprendimento-umano sia apprendimento-macchina?".

1. Sistemi intelligenti di tutoraggio: imparare attraverso l'interazione cognitiva uomo-macchina

Fin da quando le macchine hanno dimostrato la possibilità di interagire con l'umano e scambiare con esso informazioni, sono stati concepiti sistemi di istruzione automatizzata in cui la macchina dava informazioni, stimoli e consegne all'allievo e questi le elaborava e le risolveva, inviando alla macchina i suoi prodotti per avere un feedback. È questa la logica delle "macchine per insegnare" (Pressey, 1960). L'intelligenza artificiale mette però a disposizione qualcosa di più: la possibilità di avere un'interazione "intelligente", che simuli le funzioni di tutoraggio che potrebbe svolgere un tutor umano.

Sleeman e Brown coniarono nel 1982 l'espressione "Intelligent Tutoring Systems" (ITS) per riferirsi a prodotti software in grado di supportare l'apprendimento fornendo servizi tutoriali in grado di dimostrare "intelligenza" in questo ruolo. Gli ITS sono ambienti di apprendimento computerizzati in grado di aiutare gli studenti ad acquisire determinate conoscenze e abilità mediante algoritmi intelligenti che adattano il comportamento del sistema alle specifiche esigenze che emergono man mano che gli studenti procedono nel loro percorso di apprendimento (Graesser et al. 2018). I loro vantaggi sono molteplici: possono fornire supporto 7 giorni su 7 per 24 ore su 24, mettono meno in soggezione il discente rispetto a un tutor umano e sono adattivi, potendo adeguarsi dinamicamente alle esigenze specifiche di chi li usa. La ricerca dimostra l'efficacia degli ITS nelle operazioni di tutoraggio individuale per domini di conoscenza ben definiti (ad esempio, matematica o fisica, si veda Kulik, Fletcher 2016; VanLehn 2011; Bloom 1984) superando in alcuni casi quella delle interazioni di aula tradizionali (Graesser et al. 2017).

Secondo Ouyang e Jiao (2021), sono tre i paradigmi principali per definire le modalità con cui un discente umano può essere supportato nel suo apprendimento da sistemi intelligenti di tutoraggio. Il primo paradigma vede l'allievo come recipiente di informazioni fornite dalla macchina, che

è l'elemento che dirige l'apprendimento. Fanno riferimento a questo paradigma i primi ITS basati su tecniche statistico-relazionali, come ad esempio l'*ACT Programming Tutor* (Anderson et al., 1990) e il programma *Stat Lady* (Shute, 1995).

Il secondo paradigma vede l'apprendimento come effetto delle interazioni attive e reciproche tra discente e ITS, con la macchina che raccoglie informazioni dal discente durante l'interazione e propone percorsi di apprendimento personalizzato, ottimizzando in modo adattivo il modello di studente che con essa interagisce (Beck et al., 1996). I discenti dialogano con la macchina e apprendono da questa interazione reciproca. Il processo è sempre guidato dalla macchina, che però adatta i suoi percorsi alle esigenze degli utenti. Esempi di sistemi che fanno riferimento a questo paradigma sono i *Dialogue-based Tutoring Systems* (DTSs) e gli *Exploratory Learning Environments* (ELEs) basati su reti bayesiane, elaborazione del linguaggio naturale e alberi decisionali di Markov (*Exploratory environment QUE*, Metzler & Martincic, 1998).

Il terzo paradigma vede l'apprendimento come un percorso che viene diretto dal discente e in cui la macchina viene utilizzata per potenziarne gli esiti. In questa visione, la macchina non gestisce percorsi formativi più o meno strutturati e adattivi come nel primo e nel secondo paradigma, ma è parte di un sistema complesso in cui entrano in gioco una pluralità di attori umani: allievi, insegnanti, istruttori, esperti di dominio (Riedl, 2019). Il processo si configura come un'interazione tra umano e macchina, dove la macchina funge da concentratore delle informazioni fornite dai vari umani con cui interagisce, fonte informativa e partner con cui dialogare sui temi oggetto di apprendimento. Il controllo dell'intero processo è nelle mani del discente, che per questo va guidato a procedere nelle giuste direzioni (Hoc, 2000; Hwang et al., 2020; Yang et al., 2021). La macchina assiste discenti e docenti in un percorso di crescita progressiva caratterizzato da un incremento del potenziale dell'intelligenza umana derivante dall'interazione sinergica con la macchina, dove sia l'umano sia la macchina contribuiscono al buon andamento del sistema facendo ciò che sanno fare meglio (es. la macchina sintetizza ed elabora dati, l'umano assegna significati e prende decisioni). Si crea quindi una forma di cognizione umano-artificiale (Hwang et al., 2020) dove la potenza del sistema sta nell'interazione, cooperazione e integrazione tra le due entità e non nelle entità separate. Esempi di sistemi di questo tipo sono i *Real-time MOOC predictive modeling* (Le et al., 2018).

È necessario sottolineare che, in tutti i tre paradigmi, un buon sistema intelligente di supporto all'apprendimento *non si può alimentare da solo con l'informazione che trova in Rete*, dato che è altissimo il rischio di costruire la propria base di conoscenza a partire da informazioni di pessima qualità. I sistemi basati sul primo paradigma ricorrono a basi di conoscenza costruite da esperti mediante apposite interfacce di programmazione e sono quindi caratterizzati da un costo elevato associato alla loro creazione e aggiornamento. I sistemi basati sul secondo paradigma apprendono da una

duplice interazione con l'ambiente: con gli esperti di dominio per acquisire la loro expertise, codificarla, immagazzinarla e renderla fruibile; con gli studenti per ricostruire il loro modo di pensare, agire, riflettere e apprendere (Ritter et al. 2014; Trincherò 2021). L'adozione di interfacce basate su modelli linguistici ha consentito alla macchina di svolgere questa interazione in linguaggio naturale (*Natural Language Processing*, D'Mello et al., 2010), sia per ricevere informazioni sia per restituire informazioni elaborate. I moderni chatbot (*ChatGPT* e *Perplexity.ai* possono essere due esempi, per rimanere su quelli di uso gratuito), sembrano offrire le caratteristiche adeguate all'implementazione del terzo paradigma: database incrementali di informazione che vengono alimentati dal dialogo in linguaggio naturale con esperti di dominio e mettono a disposizione dei novizi basi di conoscenza di cui possono fruire con interazione in linguaggio naturale.

2. Apprendere dialogando con un tutor “intelligente”

Focalizziamoci sul secondo paradigma. Come è possibile far interagire in modo sinergico umano e macchina allo scopo di mettere in atto percorsi efficaci di apprendimento/insegnamento diretti dalla macchina? Anzitutto è necessario partire da alcuni fondamentali. Per un umano, apprendere significa *costruire buone rappresentazioni mentali* a partire dalle informazioni esperite e utilizzare queste rappresentazioni per risolvere problemi di varia natura in una pluralità di contesti. La costruzione di buone rappresentazioni mentali deve seguire alcuni principi precisi (Trincherò, 2022):

a) *Coerenza*. Per essere buone, le rappresentazioni mentali che lo studente costruisce devono essere internamente coerenti e coerenti con lo stato dell'arte del sapere in quel determinato ambito. L'interazione con una macchina che contiene una vasta base di conoscenza “esperta” e dialoga con il discente su argomenti focalizzati, può aiutarlo a scoprire lacune e fonti di incoerenza nel proprio sapere. Sul versante opposto, dialogando con la macchina, il discente può trovare fonti di incoerenza tra i saperi immagazzinati dalla macchina stessa e retroagire sulla macchina per migliorarne il database.

b) *Ripetizione*. Le rappresentazioni mentali diventano più solide se si alimentano di stimoli non occasionali ma ripetuti. Instaurando con la macchina forme ripetute di interazione dialogica all'interno di un percorso didattico continuativo che porta il discente a tornare più volte nel tempo sugli stessi contenuti, questi vengono fissati e rafforzati.

c) *Significazione*. I modelli linguistici alla base del funzionamento dei chatbot contemporanei prendono in considerazione *significanti* (ossia termini e strutture linguistiche) e non *significati* (i concetti e le idee che essi rappresentano). La significazione è un'operazione cognitiva umana che descrive la relazione tra un significante e un significato e consiste

nell'associare i simboli elaborati e prodotti dalla macchina (i significanti) ai significati che questi hanno per un umano, connettendoli secondo relazioni di senso alle proprie rappresentazioni preesistenti. La macchina può elaborare efficacemente simboli e mettere a disposizione dell'umano sintesi informative strutturate a cui poi quest'ultimo potrà assegnare più agevolmente significato.

d) *Interpretazione*. L'operazione di attribuzione di significato non è univoca e lineare, ma soggiace a margini interpretativi da parte dell'umano. Le interpretazioni possono variare a seconda delle rappresentazioni mentali pregresse e del contesto culturale in cui avviene la significazione e portano a vedere significati differenti a partire dalle stesse informazioni. Il confronto intersoggettivo tra umani supportato dalla macchina (ad esempio nel caso di più esperti che alimentano la stessa base di conoscenza) può evidenziare le diverse possibilità interpretative a partire dagli stessi stimoli e consente quindi sia di mettere alla prova le proprie interpretazioni sia di acquisirne un ventaglio più ampio e variegato.

e) *Progressione*. Buone rappresentazioni mentali possono essere affinate e ampliate assimilando in modo graduale informazioni di difficoltà progressivamente più elevata. Si parte dall'acquisire padronanza di elementi più semplici e basilari per giungere via via alla padronanza di elementi più complessi che utilizzano i primi come "mattoni" del sapere. Interagire con la macchina in un percorso didattico strutturato che propone la giusta sequenza di saperi aiuta ad acquisire con sicurezza "prima" gli elementi di sapere necessari per acquisire con sicurezza gli elementi di sapere che verranno "dopo".

f) *Riorganizzazione*. Le rappresentazioni mentali diventano migliori non solo per accrescimento, ma anche per riorganizzazione. L'interazione con la macchina e il confronto con rappresentazioni diverse che la macchina rende disponibili può aiutare a scoprire i punti deboli delle proprie rappresentazioni e offrire spunti ed esempi per riorganizzare in maniera maggiormente efficace i propri saperi.

g) *Ricostruzione*. Le rappresentazioni mentali si rafforzano se l'umano è chiamato a recuperare dalla sua memoria elementi di sapere per ricostruire informazioni mancanti, collegamenti non esplicitati, compiere inferenze. La macchina può aiutare l'umano in questo processo ponendogli domande che lo inducono a ricostruire i saperi posseduti su un dato argomento, portandolo così a identificare carenze e punti non chiari.

h) *Automatizzazione*. Le rappresentazioni mentali sono più efficaci quando vengono automatizzate. Per un discente umano questo significa usare i suoi saperi con fluidità e sicurezza, estraendo dalla sua memoria ciò che serve, quando serve e usandolo nel modo in cui serve, senza passare attraverso il controllo cosciente. Il confronto e la messa alla prova continua da parte della macchina può aiutare l'automatizzazione dell'uso delle proprie rappresentazioni in situazioni paradigmatiche definite.

i) *Feedback*. La messa alla prova delle rappresentazioni mentali ne può evidenziare punti di forza e punti di debolezza, ma anche focalizzare i

marginari di miglioramento. La macchina può dare un feedback personalizzato al discente fornendogli l'informazione che serve proprio quando serve e colmare quindi le carenze emerse.

j) *Focalizzazione*. Per poter rispondere ad esigenze mirate le rappresentazioni mentali devono essere focalizzate. Avere troppa informazione non aiuta nella focalizzazione. La macchina può aiutare il discente a gestire grossi carichi informativi facendo selezioni e categorizzazioni e aiutando quindi nell'individuare l'informazione che serve, quando serve.

k) *Modellamento*. Buone rappresentazioni vengono costruite a partire da buoni esempi di ragionamento e di azione. "Far vedere come si ragiona e come si agisce" è uno dei modi più efficaci per far apprendere una conoscenza procedurale. La macchina può aiutare a trovare esempi pratici e operativi che aiutano il discente nella costruzione di rappresentazioni efficaci.

l) *Elaborazione*. Buone rappresentazioni derivano da buone elaborazioni delle informazioni a disposizione. Buone elaborazioni implicano l'uso di una pluralità di processi cognitivi: un'informazione non va solo ricordata, ma va compresa, applicata, analizzata, valutata, ricreata, per poter dire che la si padroneggia realmente e che è entrata a far parte in modo stabile di una nostra rappresentazione mentale. In questo un ruolo chiave viene svolto dalle consegne assegnate dalla macchina al discente: consegne che chiedono all'allievo di ricordare elementi di sapere inducono un'elaborazione superficiale dell'informazione a disposizione, consegne che richiedono all'allievo di esercitare processi cognitivi di alto livello, supportati dalla macchina, favoriscono un'elaborazione maggiormente profonda e significativa dell'informazione.

m) *Recupero*. Le rappresentazioni mentali si conservano stabili nel tempo se vengono periodicamente recuperate dalla memoria e "rinfrescate", attraverso l'esercizio di una pluralità di processi cognitivi (vedere punto precedente). Recuperare frequentemente informazioni dalla memoria migliora il ricordo delle stesse. Una macchina che implementa un percorso formativo che mette frequentemente alla prova l'allievo e gli chiede di recuperare dalla memoria quanto appreso precedentemente e utilizzarlo per soddisfare consegne mirate lo aiuta a consolidarlo.

n) *Contestualizzazione*. Le rappresentazioni mentali sono più stabili se insieme ai saperi contengono i contesti di uso di quei saperi. Un concetto astratto viene meglio ricordato e meglio compreso se connesso a situazioni di vita dei discenti, che a esso danno un senso e una connessione con la realtà. La macchina può suggerire al discente contesti d'uso dei saperi e problemi reali in cui questi trovano applicazione, che danno senso e ricchezza di significato a quanto appreso.

m) *Autoregolazione*. Il discente deve essere consapevole delle proprie rappresentazioni mentali, dei loro punti di forza e dei loro limiti. Dove i limiti emergono deve agire per colmarli. La macchina può stimolare il discente nel porsi domande sulla bontà delle proprie rappresentazioni e

aiutarlo nel trovare modalità e strategie di miglioramento, avviando e supportando percorsi di riflessione metacognitiva.

o) *Emozione*. Le rappresentazioni che incorporano elementi emotivi sono più efficaci perché promuovono una più ricca attivazione cognitiva. Sia le emozioni positive sia le emozioni negative incidono sui ricordi. Le emozioni positive hanno però anche un impatto positivo sull'interesse, sul piacere nell'apprendere, sull'impegno che il discente è disponibile a approfondire nel processo di apprendimento. La macchina può aiutare a mettere in atto percorsi di apprendimento fluidi, non frustranti e a costruire materiali ed esperienze di apprendimento piacevoli e motivanti, anche in relazione alla sua capacità di mettere in atto dialoghi stimolanti con il discente.

p) *Differenze individuali*. Soggetti diversi costruiscono, a partire dagli stessi stimoli, rappresentazioni diverse in tempi diversi. Questo è legato alle diverse rappresentazioni mentali di partenza che a loro volta derivano da esperienze differenti, stimoli differenti, assegnazioni di significato differenti. La macchina, come sistema adattivo, può venire incontro alle differenze individuali mediando gli stimoli in modo da adeguarli alle preconcoscenze, alle possibilità, ai tempi soggettivi dei discenti.

q) *Intenzionalità*. Costruire buone rappresentazioni mentali su un tema dipende anche dal tempo e dalle risorse personali che il discente è disposto ad investire su quel tema. Ciascuno ha le proprie preferenze e compie le proprie scelte. La macchina può aiutare il discente a identificare i punti di contatto tra le discipline di studio e i propri interessi personali, aumentando così la significatività dell'argomento di studio per il discente.

I principi elencati caratterizzano un processo dialogico tra umano che apprende e macchina che ne supporta "intelligentemente" l'apprendimento secondo un tragitto flessibile, proponendo contenuti, attivando forme di dialogo e di valutazione formativa, guidando il discente lungo un percorso di apprendimento organizzato dagli *instructional designer* che hanno progettato l'ITS stesso (Memarian, Doleck, 2024). La macchina implementa il percorso rappresentando una sorta di coach che guida l'umano attraverso sequenze di attività predefinite che vengono via via adattate dalla macchina stessa alle esigenze del discente. Questo paradigma offre sicuramente il vantaggio di garantire ai discenti una buona guida istruttiva nell'esplorare i saperi oggetto di apprendimento, ma lo svantaggio è il costo elevato sia per l'implementazione sia per l'aggiornamento del sistema.

3. Apprendere interrogando criticamente un chatbot

Focalizziamoci ora sul terzo paradigma. L'idea di base è quella di avere un discente che impara interagendo con la macchina ma che sia in grado da solo di pianificare, orchestrare e orientare il proprio

apprendimento in direzioni precise. Questo può essere vero nell'apprendimento adulto, dove i soggetti hanno probabilmente motivazione, senso di responsabilità e un bagaglio di preconoscenze tali da poter affrontare un percorso di interazione con un sistema di intelligenza artificiale in maniera complessivamente autonoma, soprattutto se già dotati di un buon livello di expertise sugli argomenti oggetto di apprendimento. Tutto ciò non vale per gli apprendimenti scolastici. Lasciato da solo di fronte a una macchina in grado di dare risposte apparentemente valide a una miriade di quesiti, la tentazione del discente medio è quella di fidarsi ciecamente delle risposte della macchina e lasciare che questa risolva i problemi al posto suo. Ma come può controllare che le risposte e le soluzioni prodotte siano effettivamente attendibili? Un siffatto controllo potrebbe metterlo in atto l'esperto, ma il novizio non ne ha gli strumenti. Il problema è rilevante perché, come abbiamo visto, l'elaborazione simbolica che porta avanti la macchina non è detto che produca un responso dotato di significato, ossia di corrispondenza con la realtà.

Una via di uscita è progettare percorsi che l'allievo dovrà mettere in atto sotto la supervisione del docente, utili a fargli sviluppare le capacità critiche necessarie per poter valutare la qualità delle risposte offerte dalla macchina e per poter incorporare solidamente queste risposte all'interno delle proprie rappresentazioni mentali. I principi di costruzione di buone rappresentazioni elencati nel paragrafo precedente non vengono in questo caso soddisfatti da un tutor automatizzato ma dal fatto che l'allievo segue un percorso progettato dal docente che lo guida nell'interazione con il chatbot (ad esempio i già citati ChatGPT e Perplexity.ai). In mancanza di tale guida istruttiva è ovviamente molto forte il rischio per il discente di sviluppare rappresentazioni distorte e inadeguate.

Come è possibile tradurre tutto questo in pratica didattica? Il docente può ad esempio porre delle consegne che gli allievi devono svolgere singolarmente o a coppie interagendo con la macchina e documentando nel dettaglio il processo di interazione che ha portato alla costruzione della risposta. La consegna può riguardare il risolvere problemi, l'analizzare testi, il costruire progetti, il formulare scenari e idee innovative, e deve prevedere la costruzione di risposte a quesiti aperti, ossia che ammettono più soluzioni o soluzioni uniche a cui però è possibile giungere tramite percorsi differenti.

Ricevuta la consegna, gli allievi possono iniziare a interagire con il chatbot in linguaggio naturale, formulando *prompt* (ossia richieste testuali finalizzate a generare un output specifico) opportuni. Grazie ai modelli linguistici in essi implementati, i chatbot sono in grado di rispondere a un numero enorme di richieste dell'utente consultando la loro base di conoscenza. La prima difficoltà per gli allievi sta nel formulare le domande "giuste" alla macchina, ossia i giusti *prompt*, in un italiano corretto e comprensibile. Supponiamo che la consegna di partenza data dall'insegnante agli allievi sia "Vi è stata in passato un'estinzione di massa del genere umano? Quali sono state le cause? Quali sono stati gli effetti?",

gli allievi devono partire da queste domande, formulare le giuste richieste al chatbot e analizzare le risposte prodotte per controllarne il grado di affidabilità e di profondità. Non è detto infatti che le risposte del chatbot siano sempre affidabili, dato che dipendono dalle fonti a partire dalle quali la macchina ha costruito la sua base di conoscenza e dall'elaborazione delle informazioni tratte da queste mediante i modelli linguistici a disposizione. Si potrebbero quindi verificare fenomeni di scarsa affidabilità delle fonti, mancanza di pertinenza, coerenza logica, esaustività, chiarezza e profondità. In tal caso l'interrogazione deve proseguire chiedendo chiarimenti, esplicitazioni e riformulazioni delle risposte, secondo un processo di interrogazione critica molto simile a un dialogo socratico. Vediamo come affrontare alcuni dei problemi citati:

a) La macchina costruisce risposte sulla base di *informazioni tratte da fonti scarsamente affidabili*, non sottoposte a controllo scientifico (es. voci di blog, siti commerciali, fake news, siti di propaganda, ...) anziché basarsi su fonti verificate (es. articoli scientifici, pubblicazioni di autori riconosciuti, documentazione messa a disposizione da centri di ricerca affidabili, ...). Per evitare questo rischio è sempre buona norma invitare il chatbot ad esplicitare le fonti informative utilizzate (ad es. con domande del tipo "Quali fonti hai utilizzato per costruire questa risposta?", "Quali autori affermano ciò che stai dicendo?") e i processi di inferenza messi in atto (es. "Come puoi dimostrare le tue affermazioni?"), inducendolo mediante il prompt a utilizzare solo fonti scientificamente validate (es. inserendo la frase "Formula una risposta sulla base di articoli scientificamente validati").

b) La macchina produce *risposte non pertinenti con le richieste* che gli sono state poste. Le richieste poste alla macchina non devono essere generiche ma precise. Il prompt deve contenere consegne non ambigue quali "Descrivi ...", "Analizza...", "Calcola ...", "Produci ...", "Riformula ...", fornendo contemporaneamente anche le caratteristiche che l'output deve avere (es. "Produci una linea dei tempo che evidenzi la successione cronologica di eventi che ha portato all'estinzione di massa degli esseri umani avvenuta 800.000 anni fa, indicando le prove a sostegno delle affermazioni che hai prodotto."). Per avere risposte pertinenti e focalizzate sulle proprie richieste è possibile fornire nel prompt il contesto a cui la richiesta si riferisce, esempi concreti delle informazioni che si desiderano o che devono essere prese in considerazione nel formulare la risposta, le parole chiave in grado di delineare con precisione l'ambito semantico a cui ci si riferisce, le regole o i vincoli che la risposta deve rispettare. Come accennato, buone risposte si ottengono formulando con precisione le richieste in un linguaggio naturale corretto e affinando la risposta con più richieste successive (tecnica del *Prompt Chaining*), che portino maggiori specificazioni e chiarimenti. Per un migliore focalizzazione è anche possibile chiedere al chatbot di formulare delle domande all'umano perché questi possa inserire nel prompt tutte le risposte che servono per focalizzare la risposta (tecnica dell'*Ask Before Answer*).

c) La macchina *sbaglia i processi di inferenza*, accostando concetti che tra di loro non hanno relazione (es. associa l'“estinzione di massa” all'“estinzione dei dinosauri”, quando invece è specificato nel problema che l'argomento che ci interessa è l'estinzione del genere umano, non di altre specie). In questo caso, quando il discente percepisce una possibile incoerenza o elementi che non sono chiari deve scavare in profondità nelle risposte della macchina, chiedendo ad esempio “Cosa vuol dire ‘estinzione del genere umano?’”, “Cosa vuol dire ‘estinzione dei dinosauri?’”. Quando si chiedono al chatbot delle definizioni, questi in genere le recupera in modo letterale dai materiali informativi di partenza e confrontando e analizzando tali definizioni mediante intelligenza umana si può capire se l'inferenza condotta è lecita o no. Si può proseguire poi con domande del tipo “Che relazione sussiste tra ‘estinzione del genere umano’ e ‘estinzione dei dinosauri?’”, controllando in modo diretto, sulla base delle definizioni appena acquisite, la plausibilità delle inferenze fatte dalla macchina. È possibile poi far notare l'errore al chatbot dicendogli con chiarezza “La tua inferenza è errata perché ... Prova a riformularla.” e vedere come riesce a produrre una nuova risposta. Qualora questo processo non porti a risultati comprensibili, la strategia è ovviamente quella di consultare autonomamente i materiali di partenza, senza la mediazione del chatbot, o chiedere aiuto all'insegnante.

d) La macchina produce una *visione parziale dell'argomento*, che non ne rende la dovuta complessità. Per ovviare a questo problema è possibile chiedere al chatbot se esistono visioni alternative a quella proposta nella prima risposta (“Esistono risposte alternative?”, “Quali altri autori hanno affrontato il problema in oggetto? Come lo hanno affrontato?”, “Quali sono gli altri punti di vista possibili sul problema?”). È anche possibile, nel prompt di partenza, specificare che nella risposta la macchina deve far riferimento a più punti di vista alternativi, se esistono, o dichiarare – giustificandolo – che non esistono.

e) La macchina produce una *trattazione che contiene punti non chiari*. In questo caso, il discente deve chiedere alla macchina di chiarificare le sue risposte e di proporre descrizioni alternative, esempi esplicativi, chiarimenti espliciti anche tratti dai contesti di uso dei concetti in questione. Va chiesto alla macchina il significato di termini ed espressioni linguistiche non noti al discente, in modo che egli possa ricostruire e fare propri i significati connessi con l'informazione offerta dalla macchina. Dove il discente debba ricavare i significati mediante i propri processi di inferenza umana può chiedere alla macchina di controllare la plausibilità degli esiti di questi processi con domande del tipo “È vero che il termine ‘...’ significa ‘...?’”.

f) La macchina produce una *trattazione superficiale e non problematizzata*. La macchina potrebbe fornire solo informazioni banali e scontate, non accompagnati da un substrato informativo profondo, oppure dare per certi saperi che invece vanno problematizzati e messi sotto esame critico. Domande da porre alla macchina per sollecitare una maggior

problematizzazione e profondità di elaborazione possono essere “Quali sono i problemi aperti su questo tema?”, “Quali sono le questioni tuttora irrisolte?”, “Quali sono le domande lasciate in sospeso?”, “Cosa non sappiamo sulla questione?”, “Da quali assunti di base deriva ciò che sappiamo sulla questione?”. Le domande poste devono consentire di esplorare in profondità la questione, non rimanendo a livello superficiale, ma riproducendo il processo logico che un umano critico metterebbe in atto per costruire una propria opinione fondata sul tema.

Come è facile vedere, in questo approccio, il processo dialogico di interrogazione, controllo, messa in discussione, approfondimento e chiarificazione delle risposte è parte integrante dell'*apprendimento* stesso, inteso sia come apprendimento umano ma anche come apprendimento-macchina, dato che l'interazione umano - “macchina che apprende” può incidere sulla base di conoscenza della macchina stessa. Interagendo in modo dialogico e critico con la macchina, il discente ne esplora possibilità e limiti, impara a formulare prompt sempre più precisi e acquisisce via via saperi sempre più affidabili e profondi.

L'insegnante valuterà il lavoro svolto dal discente sulla base:

1) della profondità delle risposte costruite dall'allievo *con* la macchina, tenendo in considerazione che senza i giusti prompt e il giusto processo dialogico, la macchina non avrebbe mai composto da sola quelle risposte;

2) sulla base di come l'allievo è stato in grado di utilizzare le risposte prodotte dalla macchina per produrre una trattazione esaustiva, affidabile e coerente;

3) sulla base di come l'allievo ha documentato il processo dialogico di costruzione delle risposte e ha dimostrato di possedere capacità critica nel condurlo.

È possibile notare che un processo di questo tipo svolto sotto la supervisione del docente incrementa sia i saperi contenutistici degli allievi sia la loro abilità nell'approcciarsi criticamente all'informazione. In particolare, il dialogo critico con la macchina aiuta il discente ad ampliare le proprie prospettive di significato, a ridurre la possibilità di misinterpretazioni e misconcezioni in quanto studiato, ad aumentare la profondità della comprensione di concetti significativi, a ridurre il sovraccarico cognitivo, delegando alla macchina le operazioni di ricerca e di sintesi delle informazioni necessarie all'apprendimento e tenendo per sé le operazioni cognitive di livello più elevato.

4. Conclusioni: il *man-with-the-machine learning*

Sia nella prospettiva del secondo paradigma di Ouyang e Jiao sia nella prospettiva del terzo, umano e macchina agiscono come un sistema unico: l'umano contribuisce al sistema con le caratteristiche e potenzialità tipicamente umane, la macchina con le caratteristiche e potenzialità tipiche

della macchina. Si genera in tal modo un sistema di *man-with-the-machine learning* che incrementa le potenzialità di entrambe le entità nell'acquire, gestire, elaborare e costruire conoscenza.

Il *man-with-the-machine learning* si realizza nel secondo paradigma nel coinvolgimento del discente nei percorsi dialogici di apprendimento implementati sulla macchina e nel terzo nel dialogo critico e costruttivo con la macchina stessa. In entrambi i casi l'obiettivo è far sì che il discente costruisca buone rappresentazioni mentali sul tema in questione, sotto la guida della macchina (secondo paradigma) o di un docente che orienta, supervisiona e controlla il processo costruttivo (terzo paradigma).

In entrambe le prospettive, alcune conclusioni sono doverose. La prima riguarda la paura legata alla sostituzione dell'intelligenza umana con forme di intelligenza artificiale. Trattandosi di forme diverse di intelligenza diverse e complementari, la macchina ha bisogno dell'umano e l'umano ha bisogno della macchina, quindi la paura che la macchina possa sostituire l'umano è del tutto infondata, a meno che ad averla non sia un umano che ragiona proprio come ragiona una macchina.

La seconda è che per sfruttare al meglio le potenzialità associate a questi sistemi, servirà una preparazione di base molto più improntata alla precisione, al senso critico, alla conoscenza approfondita dei processi cognitivi umani e dei processi elaborativi della macchina. Si apriranno quindi nuove opportunità per nuove tipologie di qualificazione (e si ridurranno per qualificazioni più basse, dove le mansioni sono simulabili in modo affidabile con l'"intelligenza" della macchina).

La terza è che l'intelligenza artificiale non renderà gli umani meno intelligenti (a meno che non vi sia qualcuno realmente disposto a rinunciare alle proprie facoltà intellettive appiattendole sulle risposte meccaniche e prive di significato date da una macchina) ma contribuirà ad esaltare capacità tipicamente ed esclusivamente umane.

La quarta è che, se saranno disponibili a tutti, questi strumenti non favoriranno solo gli studenti che partono avvantaggiati (ossia che crescono in famiglie con un buon livello socio-culturale, che hanno disponibilità di artefatti culturali e possibilità di compiere esperienze formative diversificate) ma offriranno opportunità anche agli studenti svantaggiati, a patto che l'educazione sappia offrire loro gli strumenti di base per avvalersene.

La quinta è che questi strumenti sono già qui. Il presente articolo è il frutto di un'intensa e proficua interazione dialogica con Perplexity.ai e ChatGPT. I significati espressi sono però da attribuirsi all'autore umano.

Riferimenti bibliografici

Anderson, J. R. (2009). *Cognitive Psychology and Its Implications (7th ed.)*. New York: Worth Publishers.

- Anderson, J. R., Boyle, C. F., Corbett, A. T., & Lewis, M. W. (1990). Cognitive modeling and intelligent tutoring. *Artificial Intelligence*, 42(1), 7–49.
- Beck, J., Stern, M., & Haugsjaa, E. (1996). Applications of AI in education. *ACM Crossroads Student Magazine*, 3, 11-15.
- Bloom, B. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13, 3-16, 1984.
- D’Mello, S. K., Graesser, A., & King, B. (2010). Toward spoken human-computer tutorial dialogues. *Human-Computer Interaction*, 25(4), 289–323.
- Graesser, A.C., Hu, X., & Sottolare, R. (2018). *Intelligent tutoring systems*. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, and P. Reimann (Eds.), *International handbook of the learning sciences*. New York: Routledge.
- Graesser, A.C., Rus, V., Hu, X. (2017). *Instruction based on tutoring*. In R.E. Mayer and P.A. Alexander (Eds.), *Handbook of Research on Learning and Instruction*. New York: Routledge Press.
- Hoc, J. M. (2000). From human–machine interaction to human–machine cooperation. *Ergonomics*, 43(7), 833–843.
- Hwang, G. J., & Tu, Y. F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics*. MDPI, vol. 9(6), pages 1-19, March.
- Hwang, G. J., Xie, H., Wah, B. W., & Gasevic, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers & Education: Artificial Intelligence*, 1, Article 100001.
- Kulik, J. A., & Fletcher, J. D. (2016). *Effectiveness of intelligent tutoring systems: A metaanalytic review* in «Review of Educational Research», 86(1), 42-78.
- Le, C. V., Pardos, Z. A., Meyer, S. D., & Thorp, R. (2018). Communication at scale in a MOOC using predictive engagement analytics, in *International conference on artificial intelligence in education* (pp. 239–252). Cham: Springer.
- McCarthy, J., Minsky, M.L., Rochester, N., Shannon, C.E. (2006), A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence – august 31, 1955, *AI Magazine*, 2006, vol. XXVII, (4).
- Memarian, B., Doleck, T. (2024). A review of assessment for learning with artificial intelligence. *Computers in Human Behavior: Artificial Humans*, Volume 2, Issue 1, 2024, 100040, ISSN 2949-8821.

- Metzler, D. P., & Martincic, C. J. (1998). *Explanatory mechanisms for intelligent tutoring systems*. In *International conference on intelligent tutoring systems* (pp. 136–145). Springer.
- Ouyang, F., Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, Volume 2, 2021.
- Popper, K. R., Eccles, J. C. (1981), *L'io e il suo cervello. Materia, coscienza e cultura*. Roma: Armando.
- Pressey S. (1960), *A Simple Apparatus which gives tests and scores and teaches*. In Lumsdane, A.A., Glaser, R. (1960) (eds). *Teaching Machines and Programmed Learning*, Washington: N.E.A.,
- Riedl, M. O. (2019). Human-centered artificial intelligence and machine learning. *Human Behavior and Emerging Technologies*, 1(1), 33–36.
- Ritter, S., Sinatra, A. M., & Fancsali, S. E. (2014). *Personalized content in intelligent tutoring systems. Design recommendations for intelligent tutoring systems*. In Sottolare, R. & Grasser, A. & Hu, X. & Goldberg, B. *Design Recommendations for Intelligent Tutoring Systems, Volume 2: Instructional Management*. Memphis: U.S. Army Research Laboratory.
- Shute, V. J. (1995). Smart: Student modeling approach for responsive tutoring. *User Modeling and User-Adapted Interaction*, 5(1), 1–44.
- Sleeman, D., Brown, J. S. (eds) (1982). *Intelligent tutoring systems*. London: Academic Press.
- Trincherò, R. (2021). *Design Intelligent Tutoring Systems with AI. Brain-based principles for learning effectiveness*. pp. 501-519. In G. Panconesi, M. Guida (eds.), *Handbook of Research on Teaching With Virtual Environments and AI*. Hershey, PA: IGI Global.
- Trincherò, R. (2022). *Penso dunque imparo. Guida al potenziamento cognitivo*. Milano: Sanoma.
- VanLehn, K. (2011). *The relative effectiveness of human tutoring, intelligent tutoring systems and other tutoring systems* in «Educational Psychologist», 46, 197-221.
- Yang, S. J., Ogata, H., Matsui, T., & Chen, N. S. (2021). Human-centered artificial intelligence in education: Seeing the invisible through the visible. *Computers & Education. Artificial Intelligence*, 2, Article 100008.