

## Intelligenza artificiale e antropologia pedagogica

### Riflessioni preliminari

Massimo Baldacci

*L'articolo prende in esame il rapporto tra intelligenza artificiale e pedagogica. A questo scopo discute tale rapporto nel quadro dell'antropologia umana e dell'antropologia pedagogica. Su questa base, vengono formulate alcune riflessioni sui limiti dell'uso dell'intelligenza artificiale nell'insegnamento.*

*The article examines the relationship between artificial intelligence and pedagogy. To this end, it discusses this relationship within the framework of human anthropology and pedagogical anthropology. On this basis, some reflections on the limits of the use of artificial intelligence in teaching are formulated.*

*Parole Chiave: intelligenza artificiale, antropologia, mente, significato, insegnamento*

*Keywords: artificial intelligence, anthropology, mind, meaning, teaching*

Lo sviluppo raggiunto dalle tecnologie dell'intelligenza artificiale sollecita la pedagogia a interrogarsi sui suoi possibili rapporti con essa. Le domande che sorgono sono molteplici e articolate su più livelli. Vi sono interrogativi che si pongono sul piano pragmatico del possibile uso dell'intelligenza artificiale in sede didattica, come tecnologia del processo d'insegnamento-apprendimento. Ma vi sono anche complesse questioni che sorgono sul piano teorico, e che spingono a ripensare alcuni fondamenti della pedagogia stessa, non necessariamente per modificarli, ma quanto meno per raggiungerne una diversa consapevolezza, resa più avvertita dal confronto con la problematica dell'intelligenza artificiale. Azzardo anche che le questioni di tipo pragmatico non dovrebbero essere affrontate senza una preliminare riflessione di ordine teorico, che cerchi di conseguire una conoscenza d'insieme delle problematiche in gioco.

Su questo piano teorico, una prima rilevante questione è quella della filosofia della mente che si assume, implicitamente o esplicitamente, in sede pedagogica. Educare implica sempre una concezione della mente dell'educando, per quanto rozza e approssimativa essa possa essere. L'intelligenza artificiale pone anche la questione della sua dif-

ferenza rispetto all'intelligenza "naturale", e quindi del carattere della mente umana. E se la nostra concezione generale della mente umana si modifica, cambia anche un presupposto fondamentale dell'educazione. D'altra parte, l'essere umano non è una mente disincarnata, o un cervello in una vasca (per citare il famoso esperimento mentale di Putnam del 1995). Dalla questione della mente si passa quasi necessariamente a quella dell'essere umano, alla questione dell'antropologia. Ogni concezione educativa implica una qualche idea dell'uomo, della natura dell'essere umano in quanto tale, e particolarmente come essere educabile. La riflessione sull'intelligenza artificiale può avere anche implicazioni antropologiche.

### 1. Paradigmi antropologici e modelli d'intelligenza

Partendo da ciò, intendiamo esplorare il rapporto tra l'intelligenza artificiale e la problematica dell'*antropologia pedagogica*, sebbene in forma parziale e preliminare. Ci fermeremo perciò alla soglia del piano pragmatico, senza affrontare organicamente la questione dell'uso dell'intelligenza artificiale in sede didattica. Tuttavia, cercheremo di porre una premessa che – senza delegittimarlo – cerca di stabilire un limite generale a tale uso.

Ovviamente, l'antropologia pedagogica – pur essendo caratterizzata da una propria specificità – ha forti rapporti con l'antropologia filosofica generale. E quest'ultima è caratterizzata da indirizzi diversi, a cui qui possiamo solo accennare. Iniziamo dagli indirizzi che non chiameremo in causa in questo contributo: quello metafisico-religioso e quello post-metafisico (che fa capo a Nietzsche e alla prospettiva della "morte di Dio").

Nel discorso che svilupperemo, muoveremo da tre paradigmi: quello naturalista, quello storico-culturale e quello del post-umano. A questi tre paradigmi sono legati tre diversi modelli di intelligenza: l'*intelligenza naturale*, rintracciabile in tutti gli animali, e in particolare nei mammiferi; l'*intelligenza culturale*, propria dell'essere umano, l'unico vero animale culturale; l'*intelligenza artificiale*, che può caratterizzare le macchine informatiche, e quindi anche le ibridazioni uomo-macchina. Vediamo sommariamente questi paradigmi.

*L'antropologia naturalista*. In questo paradigma l'uomo è visto da una prospettiva biologica<sup>1</sup>, e quindi la *natura umana* (vista come

<sup>1</sup> Per il rapporto tra pedagogia e biologia si veda E. Frauenfelder, *Pedagogia e*

l'essenza biologica dell'uomo) è indagata secondo un'ottica *evoluzionista* che muove dall'*eredità animale* e dalla storia naturale del processo di *ominazione*<sup>2</sup>. In questo quadro, la versione “forte” (intenzionalmente provocatoria verso le antropologie spiritualiste) enfatizza il carattere dell'essere umano come specie animale, sottolineandone – in particolare – la stretta parentela con le scimmie antropomorfe. L'uomo viene così presentato come una “scimmia nuda”<sup>3</sup>, da rendere oggetto di uno studio zoologico teso a individuare gli schemi arcaici d'azione e reazione che determinano le forme basilari del suo agire.

*L'antropologia storico-culturale*. In questo paradigma, l'uomo è visto essenzialmente come un essere storico e culturale. In quanto essere *culturale*, la sua umanità prende forma entro i modi di vita dei differenti gruppi sociali, e negli universi di simboli e di significati condivisi e oggetto della trasmissione culturale. Inoltre, l'umanità dell'uomo si definisce nel corso del divenire storico e in rapporto ad esso, manifestando così una relatività di carattere anti-essenzialista: l'uomo è come fa sé stesso nel processo storico; ovvero, gli uomini si fanno reciprocamente tali attraverso i loro rapporti sociali. A questo proposito, la versione più spiccata è senza dubbio quella che si richiama alla lezione di Marx e alla tradizione del marxismo<sup>4</sup>.

*L'antropologia post-umana* (o *trans-umana*). Questo paradigma è indirizzato a una sorta di superamento dell'uomo, preannunciato dagli odierni fenomeni di convergenza tra la vita artificiale e le biotecnologie, che sembra destinata a riconfigurare la frontiera tra “artificiale” e “naturale”, tra *uomo* e *macchina*<sup>5</sup>. Per esempio, va in questa direzione il progetto *Neuralink*, promosso dall'imprenditore Elon Musk, che prevede l'impianto di sensori nel cervello umano, così da connetterlo direttamente a dispositivi di intelligenza artificiale<sup>6</sup>. In modo meno ra-

*biologia*, Liguori, Napoli, 1994 e F. Frabboni, F. Pinto Minerva, *Manuale di pedagogia generale*, Laterza, Roma-Bari, 2001.

<sup>2</sup> Cfr. T. Pievani, *Introduzione alla filosofia della biologia*, Roma-Bari, Laterza, 2010; I. Eibl-Eibesfeldt, *Etologia umana. Le basi biologiche e culturali del comportamento*, Torino, Bollati Boringhieri, 2001; K. Lorenz, *L'altra faccia dello specchio. Per una storia naturale della conoscenza*, Adelphi, Milano, 1973.

<sup>3</sup> Cfr. il suo *La scimmia nuda. Studio zoologico dell'animale uomo*, Milano, Bompiani, 1968.

<sup>4</sup> K. Marx, *Tesi su Feuerbach*, in U. Cerroni (a cura di), *Il pensiero di Marx*, Roma, Editori Riuniti, 1975; E. Fromm (a cura di), *L'umanesimo socialista*, Milano, Rizzoli, 1975.

<sup>5</sup> F. Pinto Minerva, F. Gallelli, *Pedagogia e post-umano*, Roma, Carocci, 2004.

<sup>6</sup> Cfr. D. Caligione, *IA istruzioni per l'uso*, il Mulino, Bologna, 2022, p. 94.

dicale, l'analogia tra mente e computer ha comunque suggerito una concezione funzionalista della mente, che la vede come una *macchina di Turing*<sup>7</sup>, ossia con un dispositivo astratto per eseguire processi di elaborazione dell'informazione che operano, su simboli formali. Ciò ha portato a suggerire un'immagine – che serba una connotazione post-umana – dell'uomo come *macchina di Turing* implementata su un hardware biologico (il corpo e il cervello).

Parallelamente possiamo concepire schematicamente tre tipi d'intelligenza<sup>8</sup>.

L'*intelligenza naturale* o animale (o, meglio, le intelligenze) è costituita dalla combinazione tra l'intelligenza della specie depositata nel suo corredo di istinti (frutto della storia della selezione naturale) e l'intelligenza individuale connessa all'adattamento a specifiche circostanze ambientali, e affidata all'apprendimento; in certe specie vi svolge un ruolo limitato anche l'imitazione sociale<sup>9</sup>.

L'*intelligenza culturale* è quella propriamente umana. Intendiamoci, nessuno nega che in quanto animale l'uomo possieda una forma di intelligenza naturale, che si rende evidente fin dai primi anni di vita come intelligenza senso-motoria<sup>10</sup>. Ma su questa intelligenza naturale sopravviene<sup>11</sup> la cultura elaborata storicamente dal genere umano, un patrimonio di credenze, significati e strumenti. Così, l'intelligenza umana si compenetra di un'attrezzatura culturale, e il principale at-

<sup>7</sup> H. Putnam, *Menti e macchine* (1960) e *Filosofia e vita mentale* (1973), in H. Putnam, *Mente, linguaggio e realtà*, Adelphi, Milano, 1993

<sup>8</sup> Forse si dovrebbe dire tre concetti diversi di intelligenza. Per parafrasare Aristotele, il termine “intelligenza” si dice in molti modi. Esso acquista un senso molto diverso a seconda che lo si riferisca agli esseri umani, agli animali o alle macchine. Quindi non si deve pensare il loro rapporto come una scala evolutiva, e dunque il rapporto tra “intelligenza umana” e “intelligenza artificiale” in modo analogo a quello tra “intelligenza animale” e “intelligenza umana”. L'intelligenza artificiale, cioè, non va vista come uno stadio successivo a quello dell'intelligenza umana. Il loro rapporto va piuttosto pensato in termini di complementarità (cfr. F. Rossi, *L'IA potrà superare l'intelligenza umana?*, in “Le Scienze”, settembre 2018). Sulla fallacia di una estensione immediata di concetti propri della vita mentale umana alle macchine “intelligenti”, si veda C. Sini, *Intelligenza artificiale e altri scritti*, Milano, Jaka Book, 2024).

<sup>9</sup> D. Mainardi, *L'intelligenza degli animali*, Milano, Cairo Editore, 2010, pp. 15-22.

<sup>10</sup> J. Piaget, *La nascita dell'intelligenza nel fanciullo*, Firenze, Giunti, 1968.

<sup>11</sup> Intendo la nozione di sopravvenienza in questi termini: A sopravviene su B se A presuppone B, ma non è riducibile a B (cfr. J. R. Searle, *La mente*, Milano, Raffaello Cortina, 2005).

trezzo è il linguaggio, legato alla capacità specificamente umana di fare significato<sup>12</sup>. Di più, l'intelligenza umana stessa è definibile come l'abilità in un campo di attività culturale<sup>13</sup>, e quindi vi sono tante forme di intelligenza quanti campi culturali<sup>14</sup>.

L'*intelligenza artificiale* (IA) consiste grosso modo nella capacità di una macchina di fare cose che richiedono intelligenza, e in particolare di comportarsi o di pensare come un essere umano<sup>15</sup>. In questo ambito si possono evidenziare due diversi programmi di ricerca. L'*IA forte* mira a creare una macchina in grado di pensare come la mente umana, ovvero di simulare i processi cognitivi e le prestazioni dell'intelligenza umana; l'*IA debole* punta invece alla realizzazione di sistemi in grado di svolgere determinati compiti in modo efficiente, così da apparire "intelligenti", ossia di emulare le prestazioni della mente umana senza pretendere di riprodurre i processi<sup>16</sup>. Il primo programma è di tipo antropomorfo, il secondo non lo è; il primo mira a simulare i processi della mente umana, il secondo soltanto a emularne le prestazioni creando sistemi utili. Circa i paradigmi di realizzazione dell'AI, la prima fase si è basata sui modelli del *computazionalismo* (che operano su simboli formali), il cui esemplare è la *Macchina di Turing* come originariamente concepita. La seconda fase ha visto un cambiamento di paradigma, con l'avvento dei modelli del *connessionismo*, che funzionano tramite le *reti neurali artificiali* (che hanno un certo grado di analogia con le reti neurali del cervello umano). I programmi computazionali muovevano da un modello teorico del processo da simulare attraverso un software; le reti neurali non richiedono invece un simile modello: si basano su potenti meccanismi statistici, e possono essere addestrate attraverso un enorme mole di dati. In questo modo, sono stati sviluppati modelli di *deep learning*<sup>17</sup> che hanno raggiunto livelli di efficienza in compiti complessi talvolta perfino supe-

<sup>12</sup> J. Bruner, *La cultura dell'educazione*, Milano, Feltrinelli, 1997.

<sup>13</sup> D. R. Olson, *Linguaggi, media e processi educativi*, Torino, Loescher, 1979.

<sup>14</sup> Cfr. anche H. Gardner, *Formae mentis. Saggio sulla pluralità dell'intelligenza*, Milano, Feltrinelli, 1987.

<sup>15</sup> Cfr. M. Fasoli, G. Piredda, *Filosofia, tecnologia e scienza della mente*, Bologna, il Mulino, 2023, p. 112.

<sup>16</sup> Cfr. Op. cit., p. 96. In questo quadro acquista rilevanza la distinzione tra "simulazione" e "emulazione" (vedi G. Fornero, *Intelligenza artificiale*, in N. Abbagnano, *Dizionario di filosofia*, Torino, Utet, 1998, p. 598).

<sup>17</sup> I. Goodfellow, Y. Bengio, A. Courville, *Deep learning*, Cambridge, Massachusetts The MIT Press 2016.

riori alle prestazioni umane, mentre i modelli computazionali avevano ottenuto risultati limitati. Di contro, i processi seguiti nell'esecuzione del compito sono trasparenti nei modelli computazionali, mentre rimangono opachi nelle reti neurali artificiali. Continueremo però ad usare la *Macchina di Turing* concepita in senso lato, come metafora di una mente artificiale capace di elaborare l'informazione.

Sulla base di questa sommaria ricognizione, possiamo azzardare alcune ulteriori analisi.

## 2. Le soglie antropologiche

L'uomo come *scimmia nuda* e l'uomo come *Macchina di Turing* sono le figure limite dell'antropologia: da una parte l'uomo ridotto a mero animale; dall'altra l'uomo trasformato in una macchina. In quanto figure estreme esse sono situate ai confini dell'umano, sulla soglia tra *l'umano e il pre-umano*, e tra *l'umano e il post-umano*. L'analisi di questi confini può perciò rivestire una funzione euristica rilevante per un'antropologia pedagogica, a partire dalla cruciale categoria dell'*educabilità*. In altre parole, un'ipotesi di lavoro interessante potrebbe essere quella di ricercare la differenza specifica che caratterizza l'antropologia umana lungo tali soglie: quella uomo/macchina e quella uomo/animale.

Iniziamo dalla *differenza tra l'uomo e la scimmia antropomorfa*<sup>18</sup>. Le ricerche biologiche hanno stabilito che l'uomo e lo scimpanzé condividono circa il 99% del loro Dna. In altre parole, dal punto di vista genetico, solo una differenza dell'1% ci separa dalle scimmie antropomorfe (sebbene, in questa frazione siano compresi moti geni regolatori, capaci di determinare effetti a cascata sulle caratteristiche fenotipiche delle due specie).

Le classiche ricerche di Köhler (1968) hanno mostrato che gli scimpanzé sono capaci di comportamento intelligente, di risolvere problemi per *insight*, a condizione che tutti gli elementi del problema siano presenti nello stesso campo spaziale-percettivo (e quindi, non se invece tali elementi sono dispersi in spazi distanti). L'assenza del linguaggio verbale, che permetterebbe di richiamare anche elementi assenti, determina dunque un limite all'intelligenza degli scimpanzé.

Inoltre, come gli uomini, le scimmie antropomorfe sono capaci di

<sup>18</sup> Sulla rilevanza pedagogica della comparazione uomo/scimmia si veda A. Visalberghi, *Problemi della ricerca pedagogica*, Firenze, La Nuova Italia, 1965.

*apprendimento sociale*, per imitazione del comportamento dei propri simili, e ciò determina la possibilità dell'esistenza di tradizioni culturali. Infatti, nei vari gruppi di scimmie antropomorfe, si possono osservare differenze intergruppo nelle abitudini condivise all'interno di ogni gruppo. Per esempio, in un gruppo di macachi, è stato osservato il costume – assente in altri gruppi – di lavare le patate prima di mangiarle<sup>19</sup>. Questo ha portato alcuni studiosi<sup>20</sup> a concludere che l'uomo non è il solo animale culturale, benché sia l'unico che si sia specializzato come tale. Tuttavia, solo nella specie umana l'apprendimento culturale ha assunto un carattere cumulativo, con la trasmissione delle scoperte culturali da una generazione all'altra, che ha permesso non solo la loro conservazione ma anche il loro progressivo incremento fino all'attuale società tecnologica. Nei gruppi di scimmie antropomorfe, sebbene siano presenti apprendimenti sociali, non si è prodotto niente del genere, anzi solitamente la tradizione ha breve durata e non sopravvive oltre la generazione che l'ha creata o la successiva.

Tomasello<sup>21</sup>, un importante primatologo, si è posto il problema del motivo per cui l'effetto “dente di arresto” (il meccanismo che impedisce alla ruota della cultura di girare all'indietro, disperdendo le innovazioni realizzate) non si sia innescato nelle scimmie antropomorfe, a dispetto del fatto che siano capaci di apprendimento sociale.

Lorenz (1973) aveva già tentato una risposta a questa domanda, asserendo che – a causa dell'assenza del linguaggio – nelle scimmie la trasmissione culturale rimaneva dipendente dal contesto e dalla presenza dell'oggetto (per esempio, le patate sopra citate). Se non si crea l'opportuna concomitanza di contesto e oggetto (del tipo, compresenza di adulto e giovane, fame e presenza delle patate) per una generazione, l'innovazione va dispersa, perché non si può avviare con la trasmissione linguistica (che è indipendente dal contesto e dalla presenza dell'oggetto).

Tomasello dà una risposta in parte diversa, e per noi più interessante. Il punto è che la trasmissione culturale informale, basata sulla mera imitazione, rimane casuale (come descritto da Lorenz, nel 1973). L'effetto “dente di arresto” può essere garantito solo dalla trasmissio-

<sup>19</sup> R. Chauvin, B. Chauvin, *Il comportamento degli animali*, Roma-Bari, Laterza, 1994.

<sup>20</sup> G. Biondi, O. Rickards, *Codice Darwin. Nuove contese sull'evoluzione dell'uomo e delle scimmie antropomorfe*, Torino, Codice, 2005.

<sup>21</sup> Cfr. *Le origini culturali della cognizione umana*, Bologna, il Mulino, 2005.

ne culturale intenzionale, attuata volontariamente da un agente culturale. Ma secondo Tomasello, nelle scimmie antropomorfe questo tipo di condotta è assente, perché difettano due condizioni che la rendono possibile.

La prima condizione è il meccanismo dell'*attenzione congiunta*. Tra i nove e i dodici mesi, il bambino passa da un modulo di attenzione diadico (adulto-bambino) a uno triadico: adulto-oggetto-bambino. Si crea, cioè, uno spazio d'attenzione condiviso tra l'adulto e il bambino e centrato sull'oggetto. L'adulto attira volontariamente l'attenzione del bambino sull'oggetto, ed entrambi prestano simultaneamente un'attenzione reciproca che include tale oggetto. Questo meccanismo di attenzione congiunta rappresenta una condizione necessaria per la trasmissione intenzionale. Ma negli scimpanzé tale meccanismo sembra assente. Anche se uno scimpanzé (diciamo giovane) osserva cosa fa un proprio simile (diciamo un adulto) con un oggetto, il secondo scimpanzé non presta attenzione al primo. Così, non si crea uno spazio d'attenzione congiunta, e perciò l'eventuale imitazione rimane spontanea, senza essere favorita da una sollecitazione d'attenzione e da una dimostrazione intenzionale. La seconda condizione, connessa alla prima, è la presenza di una *teoria ingenua della mente*. Verso i novedodici mesi, il bambino inizia a intuire la similarità dei propri conspecifici con sé stesso (gli altri sono come me). Così, il bambino giunge ad attribuire ai propri simili stati mentali analoghi ai propri: credenze, desideri, intenzioni. Egli elabora, cioè, una *teoria ingenua della mente* degli altri. Il punto è che l'attivazione di una trasmissione culturale intenzionale presuppone proprio l'attribuzione di una mente al soggetto da inculturare e una teoria della sua mente che permetta di concepire come articolare l'insegnamento (su questo concorda anche Bruner, 1997). Secondo Tomasello, gli scimpanzé non possiedono una teoria della mente dei propri conspecifici e probabilmente non attribuiscono loro neppure il possesso di una mente. Pertanto, non sono capaci di assumere l'atteggiamento della trasmissione intenzionale. Detto diversamente, negli scimpanzé è presente l'apprendimento sociale, ma non il comportamento sociale dell'insegnamento intenzionale.

Sembra quindi che solo l'uomo sia capace della trasmissione culturale intenzionale che ha garantito l'effetto "dente d'arresto" allo sviluppo della cultura. Se l'uomo è una scimmia, è però l'unica scimmia insegnante (lo riconosce anche Morris nel suo già citato lavoro). Ma questo significa anche legittimare le categorie di un'antropologia pe-



dagogica: solo in rapporto all'essere umano hanno senso categorie come quelle di educatore, educando e relazione educativa. Infatti, solo l'uomo educa i propri simili, insegna loro intenzionalmente (e quindi si fa *educatore*); e quindi solo l'uomo vede l'altro uomo come soggetto da educare (come *educando*). Pertanto, soltanto tra gli esseri umani si crea quello specifico tipo di relazione che si può denominare come *relazione educativa*.

Passiamo alla *differenza tra la mente umana e la Macchina di Turing*. A questo proposito, ci riferiremo innanzitutto alla posizione di Searle<sup>22</sup>: contrariamente alla tesi "forte" dell'*Intelligenza artificiale*, la mente umana è irriducibile a una *macchina di Turing*. Per giustificare questa posizione, lo studioso americano propone il famoso esperimento mentale della *Stanza cinese*. Una *Macchina di Turing* esegue procedure computazionali su *simboli formali* secondo una matrice che specifica le operazioni da compiere, ricevendo simboli di input e restituendo simboli di output. In questo modo, per esempio, può rispondere a domande poste da un individuo umano. Ora, supponiamo che l'individuo che pone domande lo faccia in remoto (inviando sms su un telefono cellulare, per esempio) e debba stabilire se le risposte che riceve sono state formulate da un essere umano o da una *macchina di Turing* (implementata in un computer, in questo caso). Si tratta del famoso *Test di Turing*. Supponiamo che l'esaminatore non sia in grado di stabilire se le risposte provengono da un essere umano o da un computer. A questo punto, si potrebbe dire che le prestazioni della *Macchina di Turing* eguagliano quelle dell'essere umano. Si può perciò asserire che la macchina può pensare? Secondo Turing e la tesi "forte" dell'*Intelligenza artificiale* la risposta è positiva. Ma per Searle questa risposta è errata, e la fonte dell'errore dipende dall'aver assunto una prospettiva in terza persona, secondo il punto di vista dell'esaminatore esterno, e quindi di fondarsi su un criterio comportamentista del tutto inadeguato. La prospettiva corretta per porre la questione è quella in *prima persona*, dal punto di vista della mente che risponde all'esaminatore. Per illustrare la questione, Searle sceglie il caso della comprensione di storie controllata attraverso domande (problema su cui lavorava l'intelligenza artificiale all'epoca in cui Searle scriveva il saggio in questione). Si può dire che la Macchina di Turing è capace di pensare nel senso di comprendere storie e di capire quindi il signifi-

<sup>22</sup> Cfr. il saggio del 1980, *Menti, cervelli e programmi*, ora in D. R. Hofstadter, D. C. Dennett, *L'io della mente*, Milano, Adelphi, 1992.

cato delle domande che le vengono rivolte? Per rispondere a questa domanda da una prospettiva in prima persona, Searle inventa l'esperimento della *Stanza cinese*, nel quale la macchina è grande come una stanza e al suo interno è presente un individuo umano (che svolge così il ruolo di una sorta di *demone di Turing*, secondo l'espressione di Hofstadter e Dennett<sup>23</sup>). Da una finestra di input, questo individuo riceve biglietti scritti in cinese (le domande). Egli non conosce il cinese, ma ha a disposizione una tabella (la matrice funzionale di Turing) che indica quale combinazione di simboli cinesi deve trascrivere su un biglietto di risposta a seconda della stringa di simboli cinesi che ha ricevuto. In questo modo, dal punto di vista dell'esaminatore cinese esterno, egli soddisfa il *Test di Turing* (1950): le risposte sono appropriate alle domande. Ma si può dire che l'individuo all'interno della stanza ha compreso il *significato* delle domande? No, perché egli non conosce il cinese, e quindi per lui i simboli cinesi sono meramente formali: differiscono solo per la forma grafica, ma sono privi di significato. E proprio qui sta il punto: i simboli formali con cui lavora la macchina non simboleggiano nulla per essa, sono unità differenziali caratterizzate da una sintassi, ma non da una semantica. La manipolazione di simboli formali secondo regole non costituisce comprensione. Pertanto, non si ha né significato, né riferimento a stati di cose esterne<sup>24</sup>. Come l'inquilino della stanza cinese, la macchina di Turing non comprende i significati<sup>25</sup>.

Pertanto, la mente umana è irriducibile a una mera *macchina di Turing*: la differenza specifica che caratterizza questa soglia è data dalla capacità umana di *comprendere il significato*.

A una conclusione simile è arrivato Bruner, criticando la psicologia cognitiva computazionale che riduce il processo mentale a una elaborazione di informazione (come fa, appunto una Macchina di Turing), sostenendo che la mente umana è caratterizzata piuttosto dalla capacità di *“fare significato”* (si torna nei pressi dell'uomo simbolico di

<sup>23</sup> D. R. Hofstadter, D. C. Dennett, *L'io della mente*, cit..

<sup>24</sup> Qui si dovrebbe rinviare alla semantica agostiniana illustrata nel *De Magistro*: se non si sa a cosa si riferisce una parola, essa è un mero suono. Più vicino a noi, Frege, nel 1892 ha definito il “significato” come un vettore semantico a due posti: il *riferimento* (l'oggetto o l'insieme degli oggetti designati dal termine) e in *senso* (il modo di presentare l'oggetto).

<sup>25</sup> Questa argomentazione di Searle è stata variamente commentata e criticata. Per alcune di tali critiche si veda il già citato saggio di Hofstadter e Dennett. Secondo noi, il discorso di Searle rimane convincente a dispetto di tali critiche.

Cassirer, dunque), e aggiungendo che il significato è irriducibile all'informazione. Quest'ultima assunzione appare però elusiva rispetto al rapporto tra informazione e significato. Uno studioso come Bateson<sup>26</sup>, per esempio, lungi dall'opporre significato e informazione, fonda il primo sulla seconda: l'informazione è una differenza che fa la differenza, per esempio una differenza nella forma di una espressione che fa una differenza di significato (per quanto riguarda l'espressione verbale, si pensi al valore differenziale dei fonemi rispetto al contenuto semantico<sup>27</sup>). A questo proposito, Searle ancora nel saggio del 1980 osserva piuttosto che l'espressione "elaborazione di informazioni" ha un senso del tutto diverso quando si riferisce a un essere umano (nella quale è sempre legata a una dimensione semantica), rispetto a quando si riferisce a una Macchina di Turing (nella quale ha solo una dimensione sintattica).

In ogni caso, le conclusioni di Bruner convergono con quelle che possiamo trarre dalla pista analitica di Searle, se la differenza specifica dell'intelligenza umana rispetto a quella artificiale è la capacità di *comprensione del significato*, allora l'educazione si definisce in rapporto a tale attività di fare e capire significati. E la categoria antropologica dell'*educabilità* si ridefinisce nei termini della *capacità di apprendere e comprendere significati*.

Il concetto di "significato" è sufficientemente astratto da coprire un'ampia gamma di fenomeni e di apprendimenti. L'uomo vive in un mondo di significati, e il significato è un potente strumento di riorganizzazione dell'esperienza<sup>28</sup>. Al punto che lo stesso percorso della formazione scolastica può essere pensato in larga misura nei termini di apprendimento di significati<sup>29</sup>. A ciò, riferendoci a Vygotskji<sup>30</sup>, possiamo aggiungere che i significati si definiscono in uno spazio interpersonale, condiviso (uno spazio intrinsecamente inerente al processo educativo, dunque), per essere poi interiorizzati in qualità di strumenti mentali. Quindi, l'educazione può essere concepita come una attività di costruzione, comunicazione, condivisione e comprensione di significati. Una dimensione al di fuori della portata delle Macchine di Tu-

<sup>26</sup> Cfr. i suoi due lavori *Verso un'ecologia della mente*, Milano, Adelphi, 1997 e *Mente e natura*, Milano, Adelphi, 1999.

<sup>27</sup> Cfr. A. Martinet, *Elementi di linguistica generale*, Roma-Bari, Laterza, 1977.

<sup>28</sup> Cfr. l'opera del 1925 di J. Dewey, *Esperienza e natura*, Milano, Mursia, 1973.

<sup>29</sup> Cfr. D. P. Ausubel, *Educazione e processi cognitivi*, Milano, Angeli, 1968.

<sup>30</sup> Cfr. *Storia dello sviluppo delle funzioni psichiche superiori e altri scritti*, Firenze, Giunti-Barbera, 1974.

ring. In ogni caso, se l'uomo è una macchina, è però l'unica macchina in grado di fare significato.

Un'altra critica alla identificazione tra la mente di un soggetto e una macchina computazionale è stata formulata da Thomas Nagel nel saggio *“Che effetto fa essere un pipistrello?”*. In estrema sintesi, l'argomento è il seguente. Si potrebbe avere una conoscenza completa della neurobiologia del pipistrello e delle regole che disciplinano i suoi processi computazionali. Ma, nonostante ciò, l'esperienza dell'essere un pipistrello ci sfuggirebbe completamente. Gli stati mentali e i processi cognitivi hanno un lato *soggettivo* che non viene colto dalla descrizione di tipo computazionale o biologica. A maggior ragione, l'esperienza soggettiva di un essere umano è irriducibile a meccanismi computazionali, e pertanto non può essere riprodotta attraverso essi.

Cerchiamo di riassumere i risultati di questa analisi.

L'*educabilità* dell'uomo viene configurata nei termini della capacità di apprendere e comprendere significati. Inoltre, tale educabilità pare costitutiva della stessa antropologia umana: l'uomo è l'essere educabile; è l'essere che educa i propri simili; ed è l'essere che deve essere educato per diventare ciò che è.

Infine, l'analisi delle soglie differenziali dell'umano suggerisce anche una definizione della stessa *educazione*, e con alcune specificazioni della stessa *istruzione*<sup>31</sup>: l'educazione consiste nella trasmissione intenzionale di significati e, parallelamente, nell'apprendimento e nella comprensione di tali significati. Detto diversamente, *l'educazione è un processo di comunicazione e condivisione di significati*, e tale condivisione è alla base della perpetuazione della comunità umana, come afferma Dewey in *Democrazia e educazione*, e del suo sviluppo culturale e civile.

### 3. Alcune implicazioni per l'uso pedagogico dell'intelligenza artificiale

Per intendere pienamente le implicazioni di questa analisi per l'uso dell'intelligenza artificiale in campo pedagogico, occorre qualche ulteriore specificazione sulla dimensione del significato e del linguaggio

<sup>31</sup> Qui assumiamo che l'istruzione costituisca un ambito dell'educazione caratterizzato dalla trasmissione di conoscenze-significati inerenti i saperi disciplinari.

verbale che ne è il veicolo principale. Come è noto, Wittgenstein<sup>32</sup> ha asserito che in molti casi il significato di una parola corrisponde all'uso che ne facciamo nel linguaggio. Ma ha anche chiarito che la prassi linguistica è una prassi sociale che innerva un'intera forma di vita, un intero mondo sociale. Così, per capire pienamente il significato di un discorso, per intendersi, occorre condividere una forma di vita. Infatti, nella traduzione di un'opera non basta una trasposizione meramente letterale, occorre che il traduttore abbia una certa familiarità con la cultura del Paese dell'autore. Un aspetto particolarmente rilevante di questa condivisione di uno sfondo culturale per garantire l'intesa nella comunicazione è quello del *senso comune*, del complesso di credenze ordinarie di un gruppo sociale. Davidson<sup>33</sup> ha chiarito che il significato e la credenza sono due dimensioni strettamente connesse. Una credenza ha la forma *credo che (x)*, dove (x) è una proposizione dotata di un certo significato: per esempio, *credo che (domani pioverà)*. Non si può dire che Tizio crede che (x), se egli non capisce il significato di (x). Secondo Davidson il rapporto vale anche in senso inverso: non si può capire il significato dei discorsi di Tizio se non si conoscono in buona parte le sue credenze, ossia il suo senso comune. Così, la piena comunicazione di significati esige un certo grado di condivisione di un dato senso comune (o almeno la sua conoscenza). L'apprendimento delle credenze di senso comune è un processo largamente informale, che si compie nel processo di socializzazione, e in gran parte durante la socializzazione primaria.

Questo ha implicazioni rilevanti per il processo educativo, e anche per l'istruzione in quanto processo di insegnamento-apprendimento. Tale processo, si diceva, è basato sulla comunicazione e la comprensione di significati, ma ciò presuppone non solo la condivisione di un linguaggio (di un codice della comunicazione), ma anche quella di un certo senso comune, di una data forma di vita culturale. Se l'insegnamento consiste nell'introdurre in un certo mondo culturale, esso presuppone però una preliminare condivisione del mondo della vita. Come aveva visto Visalberghi, nel suo saggio del 1990 *Insegnare ed apprendere*, la possibilità di insegnamento formale ha come propria condizione di essere stato preceduto da un ampio apprendimento informale. In questo caso, occorre che vi sia stata una vasta, anche se disorganica e in parte meccanica, educazione informale che ha condotto il

<sup>32</sup> Cfr. L. Wittgenstein, *Ricerche filosofiche*, Torino, Einaudi, 1983, § 43.

<sup>33</sup> *Verità e interpretazione*, Bologna, il Mulino, 1994, p. 205.

bambino ad assimilare il senso comune del proprio mondo sociale, prima che sia possibile intervenire con un insegnamento formale. Allora, tale insegnamento, muovendo da un senso comune e da un linguaggio almeno in parte condiviso, potrà introdurre il discente a nuovi mondi culturali e a nuovi linguaggi. La condivisione del senso comune consente all'insegnante di interpretare il contesto e le aspettative dei discenti, e quindi di regolare il proprio agire didattico. Nell'istruzione ciò avviene ordinariamente, ma come sappiamo compaiono difficoltà quando il mondo culturale del docente e quello del discente sono molto distanti (classi sociali molto diverse, dialetti, culture diverse). Allora occorre un processo di familiarizzazione reciproca tra il docente e il discente, tra i diversi mondi che hanno interiorizzato. In altre parole, nella propria teoria della mente del discente (che, come abbiamo visto, è un presupposto necessario dell'insegnamento), il docente deve incorporare il senso comune di questi (tenere conto del suo complesso di credenze), i suoi modi di fare significato. Di più, il docente deve essere capace di un accesso empatico agli stati mentali del discente, cercare di comprendere i suoi stati soggettivi (“che effetto gli fa” – per riprendere Nagel – una certa situazione o fase del processo didattico), per decidere quando e come aggiungere alla comunicazione di contenuto, quella di relazione in maniera esplicita (per esempio, incoraggiando il discente, oppure contenendolo).

Nelle prospettive di uso pedagogico dell'intelligenza artificiale si deve tenere conto di questo carattere psico-culturale complesso del processo di insegnamento-apprendimento. Come si è visto, l'insegnamento è un processo propriamente umano, che non si presenta negli animali. E qui vengo alla tesi fondamentale che intendo sostenere. Una macchina capace di simulare effettivamente il processo di insegnamento umano, nella sua complessità psico-culturale, dovrebbe necessariamente essere del tipo *IA forte*. Dovrebbe essere un'intelligenza artificiale con capacità cognitive e comprensione della soggettività molto simile a quelle umane. Oltre a una comprensione profonda dei significati scambiati, dovrebbe possedere una capacità empatica verso gli stati soggettivi del discente, necessaria in quanto l'insegnamento implica anche una relazione. Dovrebbe anche comprendere il senso comune del discente, e quindi dovrebbe a sua volta essere dotata di un senso comune. Il suo addestramento dovrebbe perciò essere in grado di simulare in modo fedele un apprendimento informale di molteplici contesti sociali ed episodi di interazione verbale e non tra gli agenti.

Posto che la ricerca riesca a realizzare forme di addestramento capaci di simulare fedelmente il senso comune, resterebbero fondati dubbi sulla capacità di fare significato nel senso forte dell'espressione (produrre e comprendere consapevolmente significati), e di attitudine empatica verso la soggettività dei discenti<sup>34</sup>. In altre parole, ritengo che l'insegnamento in senso forte rimanga una prerogativa dell'essere umano.

Forme di *IA debole* possono emulare alcuni compiti dell'insegnamento: possono fornire spiegazioni; le possono riadattare sulla base delle risposte del discente; possono formulare esercizi e test; rispondere a domande. Ma non possono valutare in modo complesso i progressi dello studente (per esempio, la sua sicurezza, autopercezione di competenza, livello di motivazione e di perseveranza ecc.). Non sono capaci di relazione empatica. Non dispongono di un senso comune (anche se ne possono emulare alcuni gradi attraverso l'addestramento su materiali tratti dalla vita quotidiana). E, infine, non hanno una capacità di comprensione autentica dei significati in gioco negli scambi relazionali docente-discente. Anche con modelli avanzati come ChatGpt, capaci di emulare forme di intelligenza conversazionale, le conversazioni sono solo il frutto di un addestramento che si basa sulla statistica e sull'enorme mole di dati. Non implicano alcuna forma di consapevolezza. I criteri che la filosofia della mente considera validi per parlare di mente sono essenzialmente due: il possesso di una coscienza, che dà una consapevolezza in prima persona dei propri stati soggettivi; e il possesso di stati mentali intenzionali, che vertono cioè su stati del mondo (quali credenze, desideri, intenzioni)<sup>35</sup>. L'*IA debole* non soddisfa questi criteri, e la realizzazione di quella forte è di là da venire. Tutto ciò non toglie che forme di *IA debole* dedicate all'insegnamento possano essere utili, e non si può quindi escludere pregiudizialmente il loro uso. Ma possono essere tanto più utili quanto più si è consapevoli dei loro limiti, e li si usa tenendo conto di questi. In parole povere, questo significa che non si deve sovraccaricare di attese l'uso didattico dell'AI. La si deve vedere soltanto come una pos-

<sup>34</sup> Boden riconosce che rispetto alla simulazione del senso comune (*L'intelligenza artificiale*, Bologna, il Mulino, 2018, p. 55), nonché delle emozioni e della empatia (*ivi*, pp. 73-5), i risultati sono – almeno fino a questo momento – molto rudimentali. I tentativi di simulare la coscienza sono invece basati su una sua riduzione funzionalista (*ivi*, pp. 121-130), che ne fa una questione molto diversa dalla coscienza fenomenica come è stata sempre intesa nella riflessione filosofica.

<sup>35</sup> Cfr. J. R. Searle *La mente*, cit.

sibile integrazione dell'insegnamento umano, che rimane e resterà sempre insostituibile<sup>36</sup>.

### Riferimenti bibliografici

- Ausubel D. P., *Educazione e processi cognitivi*, Milano, Angeli, 1968
- Baldacci M. (a cura di), *L'educazione come apprendimento del significato*, Milano, FrancoAngeli, 2008
- Baldacci M., *Riflessioni sull'antropologia pedagogica*, in "Rassegna di pedagogia", n. 3-4, 2018
- Bateson G., *Verso un'ecologia della mente*, Milano, Adelphi, 1997
- Bateson G., *Mente e natura*, Milano, Adelphi, 1999
- Biondi G., Rickards O., *Codice Darwin. Nuove contese sull'evoluzione dell'uomo e delle scimmie antropomorfe*, Torino, Codice, 2005
- Boden M.A., *L'intelligenza artificiale*, Bologna, il Mulino, 2018
- Bruner J., *La cultura dell'educazione*, Milano, Feltrinelli, 1997
- Burguier A., *Antropologia storica*, in J. Le Goff (a cura di), *La nuova storia*, Milano, Mondadori, 1980
- Caligione D., *IA istruzioni per l'uso*, Bologna, il Mulino, 2022
- Cassirer E. (1944), *Saggio sull'uomo*, Roma, Armando, 1986
- Cavalli Sforza L.L., *L'evoluzione della cultura*, Milano, Codice, 2004
- Chauvin R., Chauvin B., *Il comportamento degli animali*, Roma-Bari, Laterza, 1994
- Chomsky N. (1998), *Linguaggio e problemi della conoscenza*, Bologna, il Mulino, 1998
- Davidson D., *Verità e interpretazione*, Bologna, il Mulino, 1994
- Dewey J. (1916), *Democrazia e educazione*, Firenze, La Nuova Italia, 2000
- Dewey J. (1925), *Esperienza e natura*, Milano, Mursia, 1973
- Eibl-Eibesfeldt I., *Etologia umana. Le basi biologiche e culturali del comportamento*, Torino, Bollati Boringhieri, 2001
- Engels F. (1875-1882/ 1925), *Dialettica della natura*, a cura di L. Lombardo Radice, Roma Editori Riuniti, 1978
- Fadini U., *Antropologia filosofica*, in Rossi P. (a cura di) (1996), *La filosofia. Le filosofie speciali*, Vol. 1, Milano, Garzanti, 1996
- Fasoli M., Piredda G., *Filosofia, tecnologia e scienza della mente*, Bologna, il Mulino, 2023
- Fornero G., *Intelligenza artificiale*, in Abbagnano N., *Dizionario di filosofia*, Torino, Utet, 1998
- Frabboni F., Pinto Minerva F., *Manuale di pedagogia generale*, Roma-Bari, Laterza, 2001

<sup>36</sup> A questo proposito, una possibile prospettiva dell'uso dell'IA in campo pedagogico è quella dell'*intelligenza aumentata* (cfr. il già citato articolo di F. Rossi), ossia dell'uso dell'IA da parte di un individuo umano (un insegnante) per potenziare le proprie capacità d'insegnamento rispetto a compiti circoscritti, col mantenimento della direzione del processo complessivo da parte dell'insegnante umano, e soprattutto con la salvaguardia della sua relazione col discente.



29 – *Intelligenza artificiale e antropologia pedagogia:  
riflessioni preliminari*

- Frauenfelder E., *Pedagogia e biologia*, Napoli, Liguori, 1994
- Fromm E. (a cura di), *L'umanesimo socialista*, Milano, Rizzoli, 1975
- Gardner H., *Formae mentis. Saggio sulla pluralità dell'intelligenza*, Milano, Feltrinelli, 1987
- Geertz C. , *Antropologia e filosofia*, Bologna, il Mulino, 2001
- Goodfellow I., Bengio Y., Courville A., *Deep learning*, Cambridge, Massachusetts The MIT Press, 2016
- Habermas J. (1958), *Antropologia*, in Preti G. (a cura di), *Filosofia*, Milano, Feltrinelli, 1966
- Hofstadter D.R., Dennett D.C., *L'io della mente*, Milano, Adelphi, 1992
- Köhler W., *L'intelligenza nelle scimmie antropoidi*, Firenze, Giunti, 1968
- Lewontin R.C., *Biologia come ideologia*, Torino, Boringhieri, 1993
- Lorenz K., *L'altra faccia dello specchio. Per una storia naturale della conoscenza*, Milano, Adelphi, 1973
- Mainardi D., *L'intelligenza degli animali*, Milano, Cairo Editore, 2010
- Martinet A. , *Elementi di linguistica generale*, Roma-Bari, Laterza, 1977
- Marx K. (1844), *Manoscritti economico-filosofici del 1844*, Torino Einaudi, 1932-1968
- Marx K. (1845-1888), *Tesi su Feuerbach*, in Cerroni U. (a cura di) (1975), *Il pensiero di Marx*, Editori Riuniti, Roma
- Marx K., Engels F. (1845), *La sacra famiglia*, Roma, Editori Riuniti, 1967
- Morris D., *La scimmia nuda. Studio zoologico dell'animale uomo*, Milano, Bompiani, 1968
- Nagel T., *Che effetto fa essere un pipistrello?*, in Id., *Questioni mortali*, Milano, Il Saggiatore, 1974
- Nietzsche F. (1883-85), *Così parlò Zarathustra*, Milano, Adelphi, 2003
- Olson D.R., *Linguaggi, media e processi educativi*, Torino, Loescher, 1979
- Parisi D., *Simulazioni. La realtà rifatta al computer*, Bologna, il Mulino, 2001
- Piaget J. , *La nascita dell'intelligenza nel fanciullo*, Firenze, Giunti, 1968
- Pinto Minerva F., Gallelli R., *Pedagogia e post-umano*, Roma, Carocci, 2004
- Putnam H. (1960), *Menti e macchine*, in Putnam H., *Mente, linguaggio e realtà*, Milano, Adelphi, 1993
- Putnam H., (1973), *Filosofia e vita mentale*, in Putnam H., *Mente, linguaggio e realtà*, Milano, Adelphi, 1993
- Putnam H., *Mente, linguaggio e realtà*, Milano, Adelphi, 1993
- Putnam H., *Realismo dal volto umano*, Bologna, il Mulino, 1995
- Rossi F., *L'IA potrà superare l'intelligenza umana?*, in "Le Scienze", settembre 2018
- Schaff A., *Il marxismo e la filosofia dell'uomo*, in E. Fromm, *L'umanesimo socialista*, Milano, Rizzoli, 1975
- Searle J.R. (1980), *Menti, cervelli e programmi*, in Hofstadter D.R., Dennett D.C., *L'io della mente*, Milano, Adelphi, 1992
- Searle J.R., *La mente*, Milano, Raffaello Cortina, 2005
- Sini C., *Intelligenza artificiale e altri scritti*, Milano, Jaka Book, 2024
- Somenzi V., Cordeschi R., *La filosofia degli automi*, Boringhieri, Torino, 1986
- Tattersal I., *Il cammino dell'uomo*, Milano, Garzanti, 1998
- Tomasello M., *Le origini culturali della cognizione umana*, Bologna, il Mulino, 2005

Turing A.M. (1950), *Macchine calcolatrici e intelligenza*, in Somenzi V., Cordeschi R., *La filosofia degli automi*, Torino, Boringhieri, 1986

Visalberghi A., *Problemi della ricerca pedagogica*, Firenze, La Nuova Italia, 1965

Visalberghi A., *Insegnare ed apprendere*, Firenze, La Nuova Italia, 1990

Vygotskij L.S., *Storia dello sviluppo delle funzioni psichiche superiori e altri scritti*, Firenze, Giunti-Barbera, 1974