

Intelligenza Artificiale, Matematica e Fisica: Hic Rhodus, Hic Salta

F. LERDA^(*)

A commosso ricordo di Maria Adelaide Sneider

RIASSUNTO – *In questo articolo viene presa in esame la cosiddetta “Ipotesi forte dell’intelligenza artificiale” che sostiene l’equivalenza di principio tra l’intelligenza artificiale stessa e l’intelligenza umana; con riferimento a tale ipotesi si effettua un’analisi critica di alcuni classici teoremi logici di Gödel e Church, mettendo in evidenza il fatto che essi implicano l’esistenza di diversi “sistemi elaborativi” con livelli sempre meno significativi di formalizzazione. L’analisi delle caratteristiche di questi sistemi porta a concludere che il formalismo non è sufficiente a cogliere l’essenza profonda della creatività matematica, ed in particolare che l’“Ipotesi forte” di cui sopra non è accettabile. La stessa conseguenza si ricava attraverso considerazioni fisiche sul funzionamento dei calcolatori elettronici.*

ABSTRACT – *This paper investigates the so called “Strong Artificial Intelligence Hypothesis” that asserts the fundamental equivalence between human and artificial intelligence. With reference to such an hypothesis a few classical logical theorems by Gödel and Church are critically analysed, pointing out the existence of a number of “information processing systems” with step by step less significant levels of formalization. The analysis of such systems leads to the conclusion that formalism is not sufficient to grasp the deep nature of the mathematical creativity, and that the above quoted “Strong hypothesis” is not acceptable. The same conclusion is obtained through physical considerations on how computers work.*

^(*)Lavoro svolto nell’ambito di progetti finanziati dal Ministero dell’Università e della Ricerca Scientifica e Tecnologica e dal Consiglio Nazionale delle Ricerche.

KEY WORDS – *Artificial intelligence - Algorithmic procedure - Formal system - Logical system - Physical symbol system - Creativity.*

A.M.S. CLASSIFICATION: 03A05 - 68T01

In [1] G.C. ROTA scrive: "An authentic concept of mathematical truth must emerge from a dispassioned examination of what mathematicians do, rather than from what mathematicians *say they do*, or from what philosophers think mathematicians ought to do". In campo epistemologico questa saggia norma è stata spesso disattesa, e lo è anche oggi di frequente nello studio dei rapporti fra intelligenza artificiale ed intelligenza umana.

L'Intelligenza Artificiale è una scienza molto giovane, sviluppatasi nella seconda metà di questo secolo; la matematica ha invece una storia plurimillenaria e la filosofia della matematica, che ne ha accompagnato l'evoluzione, è passata attraverso vari tentativi e varie ipotesi, intrecciandosi in modo particolare con le ricerche tecniche sui fondamenti della disciplina. Sono nate "scuole" spesso in contrasto fra di loro (intuizionismo, logismo, formalismo, sperimentalismo,...) nessuna delle quali ha saputo cogliere in pieno l'essenza del pensiero matematico. La lettura degli scritti dei grandi matematici che si sono occupati di questi problemi (si veda, ad esempio, lo splendido saggio [2] di HADAMARD) fornisce la misura della estrema complessità dei problemi stessi e della mancanza di certezze in molti punti fondamentali; tale lettura prova anche la serietà, la prudenza, in genere l'umiltà di quegli autori nell'affrontare le problematiche in questione.

È invece impressionante vedere con quanta sicurezza molti scienziati informatici parlano oggi dell'"intelligenza", ad esempio formulando quella che è chiamata "Ipotesi Forte dell'Intelligenza Artificiale" ("Strong Artificial Intelligence Hypothesis", S.A.I.M.), che suona così: "il computer, opportunamente organizzato, possiede le condizioni necessarie e sufficienti per la generale azione intelligente", intendendo quest'ultima nel modo più ampio possibile, includente intuizione, sentimenti, creatività, ecc. Chiameremo questa posizione *positivismo informatico*.

Le obiezioni a tale concezione massimal-positivistica sono molteplici, come ho avuto occasione di precisare in [3] e [4].

Tra l'altro, tutti i precedenti storici di quello che abbiamo appena

chiamato *positivismo informatico* (determinismo meccanicistico di Laplace, "programmi matematico-logici" di Russell, Brouwer, Hilbert, positivismo logico del Circolo di Vienna, ecc.) sono falliti. Andando oltre, ritengo possibile affrontare *in termini scientifici*, sia dal punto di vista matematico e logico sia in prospettiva fisica, il problema dei rapporti fra intelligenza artificiale ed intelligenza umana, e mi propongo in questo articolo di fornire un primo schema delle linee generali di tale ricerca.

Prendiamo in esame tre classici teoremi, due di Gödel ed uno di Church, relativi questioni di indecidibilità per sistemi logici.

In un'analisi fatta in [5] Rosser riconosce in questi teoremi la presenza di due logiche: Una "logica del discorso ordinario", M , con cui vengono realizzate le dimostrazioni (metalinguaggio) ed una logica L cui si riferiscono i teoremi stessi. Il sistema logico L è un sistema formale; ciò significa tra l'altro che le proposizioni di L sono delle formule aventi una ben definita struttura propria di L stesso. Queste proposizioni ammettono interpretazioni quali affermazioni del discorso ordinario, affermazioni che possono risultare vere o false. Non tutte le affermazioni del discorso ordinario trovano la loro espressione in proposizioni di L . Tra i simboli di L compare il seguente: \sim avente significato di negazione. Nel sistema L è definito un processo secondo il quale si possono dimostrare delle proposizioni di L partendo da certe proposizioni di partenza (gli assiomi); questo procedimento è puramente formale, nel senso che non fa riferimento al significato, cioè all'interpretazione delle proposizioni di L .

Va osservato che le caratteristiche elaborative intrinseche della creatura umana son tali per cui qualsiasi sistema formale trattato dall'uomo risulta *generato in termini finiti*: i suoi assiomi possono essere in numero infinito, ma sono in ogni caso specificati con *descrizioni metalinguistiche finite*; ogni proposizione di L è finita (contiene un numero finito di simboli dell'alfabeto di L), ogni dimostrazione di L è finita. Se l'alfabeto di L contiene infiniti simboli distinti, essi sono specificati con una descrizione metalinguistica finita. E così via. L'uomo non può, per sua natura, generare sistemi formali che non rispettino tali vincoli. Il sistema formale L è detto *semplicemente coerente*⁽¹⁾ se non esiste alcuna proposizione P di L tale che sia P sia $\sim P$ risultino deducibili in L . Per quanto possa

⁽¹⁾Contrariamente all'uso corrente, adotto il termine "coerente" come corrispondente italiano dell'inglese "consistent" e non il termine "consistente" che ritengo linguisticamente scorretto.

apparire strano, si dimostra che in un sistema formale F semplicemente coerente possono esistere proposizioni dimostrabili che esprimono, in loro interpretazioni, affermazioni false.

A Tarki ha costruito un sistema formale semplicemente coerente in cui, scegliendo opportunamente la proprietà Pr , si possono dimostrare tutte le proposizioni del sistema che esprimono le seguenti affermazioni (in numero infinito) nell'interpretazione in termini di numeri naturali.

Non tutti i numeri naturali hanno la proprietà Pr

1 ha la proprietà Pr

2 ha la proprietà Pr

3 ha la proprietà Pr

.....

n ha la proprietà Pr

.....

Un sistema logico si dice ω -coerente se per nessuna proprietà Pr si verifica il fatto appena citato. La ω -coerenza implica la semplice coerenza ma non viceversa.

Dopo queste premesse, non rigorose in senso assoluto né complete, ma sufficienti per i nostri scopi, possiamo enunciare i teoremi di cui sopra:

PRIMO TEOREMA DI GÖDEL. *Per ogni sistema L di una classe estremamente ampia di sistemi logici esistono proposizioni P tali che né P né $\sim P$ risultano dimostrabili;*

SECONDO TEOREMA DI GÖDEL. *Per ogni sistema L di una classe estremamente ampia di sistemi logici, la semplice coerenza di L non può essere dimostrata in L ;*

TEOREMA DI CHURCH. *Per ogni sistema L di una classe estremamente ampia di sistemi logici non esiste alcun procedimento effettivo, cioè nessun algoritmo per individuare quali proposizioni di L siano dimostrabili.*

Le dimostrazioni di questi teoremi assumono a volte l'ipotesi della ω -coerenza, altre volte quella della semplice coerenza. Quello che è importante notare è che i sistemi per cui valgono i succitati teoremi comprendono molti fra i più interessanti sistemi formali logici e matematici, ad esempio il sistema N dell'aritmetica del primo ordine di Peano.

Come conseguenza dei risultati di Gödel e Church si ha che le due classi delle proposizioni dimostrabili in L e delle proposizioni che esprimono affermazioni vere in interpretazioni di L non risultano necessariamente coincidenti. Ma la conseguenza che a molti matematici è apparsa più sconvolgente è stata la dimostrazione della non realizzabilità, in linea di principio, del celebre *programma di Hilbert*.

Tale programma era stato formulato nel 1917 dal grande matematico tedesco e suonava così:

- 1 - realizzare una formalizzazione rigorosa dei vari campi della matematica e della logica;
- 2 - dimostrare, con metodi finitistici, aventi un livello di evidenza tale che la validità delle loro conclusioni non ammetta dubbi, la *coerenza*, la *completezza* e la *deducibilità* di tali formalizzazioni.

I teoremi di Gödel e Church di cui sopra ed altri ad essi connessi portano alla conclusione che il programma di Hilbert risulta, proprio in linea di principio, inattuabile.

Questo fatto si presta a diverse considerazioni ed apre la via, a mio avviso, ad interessanti prospettive di ricerca.

Cominciamo con l'osservare come l'analisi dei teoremi di Gödel e Church di cui abbiamo sopra parlato, e più in generale l'analisi dei sistemi logici e matematici, delle loro definizioni, delle dimostrazioni dei loro teoremi, porti a riconoscere l'esistenza di non due logiche soltanto: la logica formale L studiata e la logica M con cui si definisce L e si conducono le dimostrazioni dei suoi teoremi; esiste infatti un "terzo sistema", quello mediante il quale l'uomo procede ad *individuare* i teoremi che vuole dimostrare ed a *scegliere* i procedimenti dimostrativi che *espone* nella lingua M . A questo proposito gli sviluppi della logica matematica nei primi decenni di questo secolo sono particolarmente illuminanti, anche se, almeno per quanto è a mia conoscenza, finora non ne siano state tratte le più profonde conseguenze. Per inciso, ciò trova forse una spiegazione, almeno parziale, in questo fatto: nella prima metà del nostro secolo molti grandi matematici si sono occupati in modo sostanziale dei fondamenti della loro disciplina e più in generale di filosofia e psicologia della matematica; ricordiamo, fra gli altri, H. Poincaré, J. Hadamard, B. Russell, H. Weyl. In seguito i matematici sono quasi completamente scomparsi da questo campo di ricerche.

Riprendendo il discorso sull'evoluzione della logica matematica, os-

serviamo intanto che *l'intelligenza umana si manifesta sostanzialmente in quello che abbiamo chiamato il "terzo sistema"*, a proposito del quale si incontrano situazioni emblematiche, su cui ci soffermiamo un momento.

Secondo le stesse parole di Gödel (v. [6] p. 18 e segg.) *la concezione oggettivistica della matematica, della metamatematica e del ragionamento transfinito* è stata fondamentale per il suo lavoro nel campo della logica. L'impostazione di pensiero di un ricercatore può rivelarsi decisiva nella ricerca. Gödel osserva che un articolo di Skolem del 1922 conteneva tutti gli elementi per dedurre, in forma quasi banale, il teorema di completezza per la logica formale, ma né Skolem né altri logici del tempo trassero quella conclusione; sempre secondo Gödel la ragione di questo fatto quasi incredibile è il rifiuto, da parte di quei logici, del ragionamento non finitario in metamatematica. Ancora, è estremamente interessante la fondamentale differenza di impostazione mentale fra Hilbert e Gödel che si rivela in modo evidente a proposito dell'ipotesi del continuo: Gödel presuppone tutti gli ordinali come dati, ed accetta definizioni non costruttive, Hilbert cerca di costruire gli ordinali e considera soltanto definizioni ricorsive. Per Gödel la posizione di Hilbert è un errore filosofico.

Ora, i procedimenti di definizione e dimostrazione relativi ad un dato sistema logico L avvengono nella lingua M ma le scelte degli "oggetti" da definire, dei teoremi da dimostrare, dell'impostazione di pensiero cui fare riferimento, appartengono al terzo livello, e costituiscono i momenti in cui più profondamente si manifesta l'intelligenza.

I risultati di Gödel (e degli altri logici che hanno lavorato nella stessa prospettiva) dimostrano che i sistemi logici formali, gli algoritmi, le macchine di Turing, i computer (opportunamente organizzati) sottostanno a specifiche restrizioni. Tale fatto è stato utilizzato da qualche autore contrario all'ipotesi Forte dell'Intelligenza Artificiale in questo modo: *Le macchine presentano ben definite limitazioni nell'elaborazione di informazioni, l'uomo non ha tali limitazioni; quindi l'uomo non è una macchina.* Questa affermazione è respinta, a ragione, dai sostenitori dell'ipotesi forte di cui sopra, perché non prova nulla. Intanto bisogna dimostrare che l'uomo non ha tali limitazioni, non basta affermarlo. Ma il problema è addirittura un altro: i risultati di Gödel specificano delle *limitazioni dei procedimenti algoritmici in quanto tali*; queste limitazioni permangono sia quando i procedimenti algoritmici vengono realizzati da macchine sia quando sono attuati dall'uomo. Le macchine cui fanno riferimento i so-

stenitori della S.A.I.H. - i "positivisti informatici" - possono, per loro natura, eseguire solo elaborazioni algoritmiche. Per confutare la posizione di questi positivisti la strada è quindi quella di trovare motivi per dimostrare, o, almeno, rendere fortemente plausibile il fatto che la mente umana trascenda il procedere algoritmico.

Nello studio dei rapporti tra intelligenza umana e macchine i risultati di Gödel sono comunque estremamente significativi. In tale contesto una loro prima interpretazione è la seguente: Hilbert credeva di aver colto la natura profonda della matematica: possibilità di formalizzazione rigorosa dei vari settori della disciplina, possibilità di dimostrazione della coerenza (o della non coerenza) di tali formalizzazioni, risolubilità in linea di principio di ogni problema matematico (nel senso che ciò che esso afferma può essere sempre dimostrato vero o falso), coincidenza delle nozioni di deducibilità e verità. I teoremi di Gödel provano che tale concezione è troppo semplicistica: la matematica è molto più complessa e non è riducibile a questa concezione puramente formalistica, esattamente come la concezione puramente "linguistica" della realtà secondo L. Wittgenstein è risultata troppo riduttiva, come hanno dimostrato inequivocabilmente le analisi critiche di Popper ed altri sulle posizioni del positivismo logico del Circolo di Vienna. Si ricordi che già Eistein aveva scritto: "Tutto dovrebbe essere reso il più semplice possibile, ma non più semplice".

In termini più generali, ci troviamo di fronte a questa situazione: il "fondamento" classico della matematica è costituito dalla logica matematica e dalla teoria degli insiemi; l'intera matematica si ottiene, in un contesto formale, definendo i suoi oggetti in termini di insiemi e deducendone i teoremi a partire dagli assiomi del sistema ZFC di Zermelo-Fraenkel con l'assioma di scelta. È comprensibile la speranza che questo processo rigoroso non generi antinomie. Purtroppo non solo non disponiamo di alcuna dimostrazione di coerenza al riguardo, ma i risultati di Gödel provano che non si può ottenere tale dimostrazione all'interno del sistema stesso della matematica così concepita. La stessa situazione si ritrova per le forme alternative di "fondazione" finora tentate, per cui *allo stato attuale non esiste alcuna organizzazione concettuale della matematica completamente sicura*. (V. anche [4]).

Possiamo concludere che il *formalismo non è sufficiente a cogliere l'essenza profonda della matematica*.

Ora, osserviamo come proprio la matematica, inclusa la logica, costi-

tuisca, fra tutti i "prodotti" dell'intelligenza umana, quello che possiede il più alto grado di specificazione e di non-ambiguità rispetto alle altre discipline: fisica, chimica, biologia, sociologia, per non parlare di arte, letteratura, ecc. Sembra quindi imporsi fortemente la seguente congettura: *la matematica e la logica, includenti sia i "risultati" ottenuti sia l'"attività" matematica e logica realizzata dall'uomo per ottenere tali risultati, non sono esauribili in termini puramente algoritmici*. Abbiamo detto per prudenza "congettura", ma la situazione è tale da lasciar pensare ad una validità scientifica di tale affermazione.

A molto maggior ragione si è portati a ritenere non esauribile in termini puramente algoritmici la *generale attività intelligente* dell'uomo, in particolare il "terzo sistema" di cui abbiamo parlato sopra.

Affrontiamo ora il problema dei rapporti fra intelligenza umana e macchine da un punto di vista completamente diverso.

Il computer, organizzato secondo quanto indicato, ad esempio, da Newell e Simon, costituisce ciò che quegli autori chiamano il *sistema fisico di simboli* (*physical symbol system*) e che, a loro giudizio (S.A.I.H.) possiede le condizioni necessarie e sufficienti per la generale azione intelligente.

Tale sistema è oggi realizzato elettronicamente, con componenti attivi (transistori) e passivi (resistenze, ecc.) ottenuti in forma integrata con tecniche di microelettronica; il linguaggio con cui opera, in termini effettivi di macchina, è un linguaggio discreto con un alfabeto costituito da soli due caratteri, 0 ed 1, fisicamente ottenuti ad esempio, da due diversi valori di tensione elettrica.

Ora, in linea di principio, tali sistemi si possono realizzare con elementi puramente meccanici, e di una meccanica particolarmente "elementare": contrappesi, ruote dentate, contatti che si aprono e chiudono, ecc.; ciò è, in realtà quello che C. Babbage progettò, e parzialmente realizzò, con il nome di "analytical engine".

Da notare che questi elementi meccanici permettono la realizzazione dell'hardware, la realizzazione dei programmi, il funzionamento della macchina nell'esecuzione dei programmi, permettono cioè la realizzazione integrale del "physical symbol system" e del suo funzionamento.

Se ora fosse vera l' "Ipotesi Forte dell'Intelligenza Artificiale", se cioè ogni azione intelligente fosse di natura algoritmica e quindi attuabile in termini di sistema fisico di simboli, *tutta l'intelligenza, in linea di princi-*

pio, si ridurrebbe alla meccanica, anzi a quel tipo "banale" di meccanica.

Ma sappiamo già che la fisica *non* è riducibile alla meccanica: l'introduzione dell'elettromagnetismo trascende la meccanica: le grandezze elettromagnetiche *non* derivano da centimetro grammo, secondo. Tutto ciò quantomeno butta molta acqua sul fuoco del positivismo informatico.

BIBLIOGRAFIA

- [1] G.C. ROTA: *The Concept of Mathematical Truth*, Convegno Internaz. "La verità nella Scienza", Accademia Nazionale dei Lincei, Roma 13-14 ottobre 1989.
- [2] J. HADAMARD: *The Psychology of Invention in the Mathematical Field*, Dover Publ., New York, 1954.
- [3] F. LERDA: *Precedenti, Premesse, Possibilità e Limiti dell'Intelligenza Artificiale*, Nuova Civiltà delle Macchine, VI, 1/2, 1988.
- [4] F. LERDA: *La Creatività in Matematica*, Conferenza tenuta al Convegno Internazionale "La Creatività nella Scienza e nell'Arte", Locarno, 6-8 ottobre 1988, Nuova Civiltà delle Macchine, VIII, 2/3, 1990, Edizione inglese in *Creativity in the Arts and Science* (W.R. Shea and A. Spadafora eds), Science History Publ., U.S.A. 1990.
- [5] M. DAVIS (ED.): *The Undecidable*, Raven Press, Hewlett, New York, 1965.
- [6] H. WANG: *Dalla matematica alla filosofia*, Boringhieri, Torino, 1984.

*Lavoro pervenuto alla redazione il 27 novembre 1990
ed accettato per la pubblicazione il 7 dicembre 1990
su parere favorevole di P. Benvenuti e di P.E. Ricci*

INDIRIZZO DELL'AUTORE:

Francesco Lerda - Dipartimento di Matematica - Università di Torino - Via Carlo Alberto, 10
- 10123 Torino - Italia