

## Alla ricerca della mente

# L'intelligenza artificiale tra filosofia e vita di tutti i giorni



“Before we work on artificial intelligence  
why don't we do something about natural stupidity?”

**Steve Polyak**

# Realtà e Rappresentazione

Una caratteristica dell'intelligenza è la possibilità di indagare la realtà.

Non c'è dubbio che ogni nostra conoscenza incomincia con l'esperienza; da che infatti la nostra facoltà conoscitiva sarebbe altrimenti stimolata al suo esercizio, se ciò non avvenisse per mezzo degli **oggetti** che colpiscono i nostri sensi, e, per un verso, danno origine da sé a **rappresentazioni**, per un altro, muovono l'attività del nostro intelletto a paragonare queste **rappresentazioni**, a riunirle o separarle, e a elaborare per tal modo la materia grezza delle **espressioni sensibili**, per giungere a quella conoscenza degli **oggetti**, che chiamasi esperienza?  
(Kant, Critica della Ragion Pura)

L'Universo non è un tutto unico ma è fatto di **cose** - **oggetti** (τὰ ὄντα).

Non è possibile indagare le **cose** senza nominarle o rappresentarle.

**Vi sono due modi di intendere la rappresentazione.**

## Convenzionale

- I nomi sono *puri purissimi accidenti* (Manzoni)
- I nomi e gli attributi si devono accomodare all'essenza delle cose e non l'essenza ai nomi perché prima furono le cose e poi i nomi (Galileo)
- Quella che chiamiamo rosa con un altro nome conserverebbe lo stesso dolce profumo (Shakespeare)

## Sostanziale

- Ἐν ἀρχῇ ἦν ὁ λόγος ... et verbum caro factum est (Giovanni)
- *Nomen omen* ovvero *Conveniunt rebus nomina saepe suis* (Riccardo da Venosa)
- Aspetti magici del nome

L'associazione dei concetti ai nomi è un problema di dimensioni epocali.

Nomi (concetti) facili:

Pera, Mela, Sole

Nomi (concetti) difficili (tecnici):

Protosincrotone, Conica, Coleottero, Angina, Bugnato

Nomi (concetti) problematici:

Pensiero, Intelligenza, Mente, Coscienza

È più complicato di quello che sembra:

- I termini vanno definiti in base ad altri termini
- Le affermazioni hanno diversi gradi di verità, falsità, problematicità

*Questa frase è falsa*

*Il libro (non) è sul tavolo*

*Dio (non) esiste*

*La Teoria della Relatività Generale e l'Evoluzione sono "fatti"*

*Con i vaccini si è sconfitto il vaiolo*

*Devo vaccinare mio figlio?*

*I cori angelici sono i seguenti (ciascuno associato ad una sfera celeste):*

*Serafini, Cherubini, Troni, Dominazioni, Virtù, Potestà, Principati, Arcangeli, Angeli*

*La terra è piatta*

*Se mia nonna avesse le ruote sarebbe un carretto*

*La velocità della luce nel vuoto  $c$  vale 299 792 458 m/s*

*La costante di Planck  $h$  vale 6,62606957015... E-34 J s, (fino al 20/5/2019)*

*La costante di Planck  $h$  vale 6,62606957015 E-34 J s, dal 20/5/2019*

*Quel che è possibile è obbligatorio*

*L'uomo è intelligente*

*L'uomo ha una coscienza*

*Le macchine sono intelligenti*

*Le macchine potranno essere intelligenti*

*Le macchine (non) potranno (mai) avere una **coscienza***

*I grossetani sono esseri inferiori*

# L'intelligenza delle macchine

Mi propongo di considerare la questione: “Possono pensare le macchine?” Si dovrebbe cominciare con il definire il significato dei termini **macchina** e **pensare** ... Se il significato delle parole **macchina** e **pensare** deve essere trovato esaminando le parole stesse attraverso il loro uso comune è difficile sfuggire alla conclusione che tale significato e la risposta alla domanda **possono pensare le macchine?** vadano ricercati in una indagine statistica.

(A. Turing, Computing machinery and intelligence)

*Di nuovo un problema di nomi*

# Medium cognitivo

Percezione: attraverso una combinazione di sensi ed esperienza

## Il mito della Caverna

Dentro una dimora sotterranea a forma di caverna, con l'entrata aperta alla luce e ampia quanto tutta la larghezza della caverna, pensa di vedere degli uomini che vi stiano dentro fin da fanciulli, incatenati gambe e collo, sì da dover restare fermi e da poter vedere soltanto in avanti, incapaci, a causa della catena, di volgere attorno il capo. Alta e lontana brilli alle loro spalle la luce d'un fuoco e tra il fuoco e i prigionieri corra rialzata una strada. Lungo questa pensa di vedere costruito un muricciolo, come quegli schermi che i burattinai pongono davanti alle persone per mostrare al di sopra di essi i burattini. Immagina di vedere uomini che portano lungo il muricciolo oggetti di ogni sorta sporgenti dal margine, e statue e altre figure di pietra e di legno, in qualunque modo lavorate ... credi che tali persone possano vedere ... altro se non le ombre proiettate dal fuoco sulla parete della caverna che sta loro di fronte? ... Se quei prigionieri potessero conversare tra loro, non credi che penserebbero di chiamare oggetti reali le loro visioni?

(Platone, Repubblica)

# Approccio “dogmatico” al pensiero

## Ego sum res cogitans

Le pecore fanno “bee”

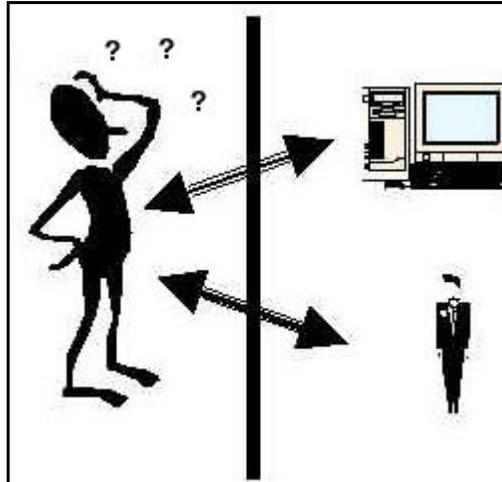
Il fuoco brucia

Gli uomini pensano

E le pecore?   Le macchine?   I grossetani?

# Approccio “comportamentista” al pensiero

## *Il test di Turing classico*



# Test di Turing implicito

## Mente? Intelligenza? Sentimento? Coscienza?

- Tutti lo facciamo, decidendo se l'interlocutore è intelligente
- I bambini parlano con i loro pupazzi
- Molti parlano con i gatti e con i cani come fossero bambini piccoli

- Alcuni parlano con le piante:

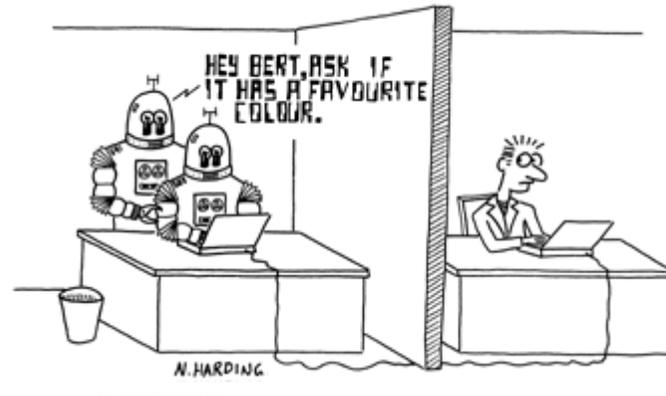
La zia Bettina parlava col dittamo



- Se qualcuno parla con il frigo o con le scarpe allora ha dei problemi, oppure è molto solo

# Test di Turing inverso

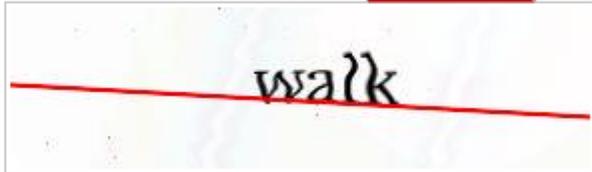
A volte le macchine ci fanno domande per sapere se **NOI** siamo umani o macchine



**Word Verification**

This step helps Yahoo! prevent automated registrations.

Enter the word as it is shown in the box below.



If you can not see this image [click here](#).

Si noti che tutto quello che si può dire per dimostrare la **presunta** superiorità dell'**uomo** sulle **macchine** si potrebbe dire (**ed è stato detto in passato**) per dimostrare una **presunta** superiorità dell'**uomo** sulla **donna** o di una razza sull'altra.

“Le **macchine** non possono pensare perché sono **diverse** dall'**Uomo**”

Nel X secolo andare sulla **Luna** era ritenuto **impossibile**, lo era di fatto, stante la tecnologia dell'epoca, ma non certo perché la **Luna** era nella **sfera degli Angeli**.

I **viaggi nel tempo** ancora oggi sono **tecnologicamente impossibili** e non è chiaro se vi saranno mai gli strumenti per farli.

Le **macchine intelligenti** sembrano un obiettivo possibile, certamente più difficile di andare sulla **Luna**, e fuori dalla portata della attuale tecnologia, ma teoricamente più plausibile dei **viaggi nel tempo**.

Per poter dire qualcosa su **macchine coscienti** bisognerebbe prima mettersi d'accordo su cosa è la **coscienza** nell'uomo

Differenza tra visione **comportamentista** e visione **dogmatica**.

# Intelligenza artificiale ?

## Approccio “storico” (problem solving)

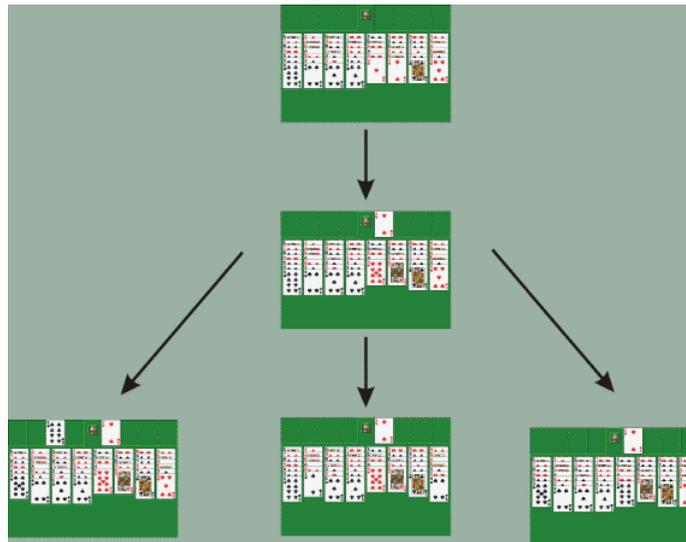
Algoritmi basati su regole

Per giocare a scacchi (Deep Blue vs. Kasparov, Stockfish)

Per tradurre da una lingua all'altra

ecc.

## Ricerca in uno spazio di stati



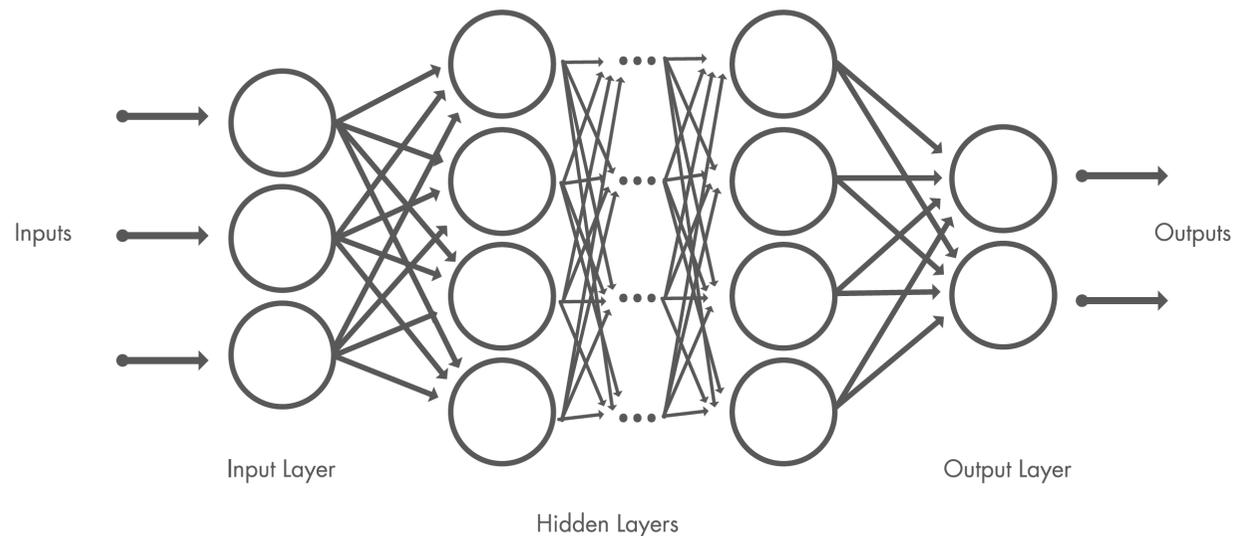
Albero del solitario

# Deep learning

## Minimizzazione di funzioni multivariate

### Algoritmi basati su apprendimento

A deep neural network combines multiple nonlinear processing layers, using simple elements operating in parallel and inspired by biological nervous systems. It consists of an input layer, several hidden layers, and an output layer. The layers are interconnected via nodes, or neurons, with each hidden layer using the output of the previous layer as its input.



Con meno regole, senza l'esperienza umana, cade completamente l'obiezione *so come funziona* e anche la critica della *stanza cinese* (Searle)

# Applicazioni

Riconoscimento immagini

**All'aeroporto di Atlanta il check in si fa mostrando il volto**



Traduzione automatica

Assistenti vocali (Siri, ecc.)

Guida automatica

Diagnosi mediche

Protein Folding

## Pittura automatica (sic)



Il ritratto di “Edmond de Belamy de la Famille de Belamy”, prodotto di un algoritmo elaborato dal collettivo francese Obvious è stata battuto all’asta da Christie’s per 432 mila dollari. Per raggiungere l’obiettivo, il software è stato alimentato con 15.000 ritratti classici dal XIV al XX secolo; e così il software ha imparato a *capire le regole del ritratto* e poi ha generato una serie di nuove immagini.

L’algoritmo è basato su un **Generatore** che crea la nuova immagine e un **Discriminatore** che fa da antagonista tentando di stabilire quale immagine è opera di un essere umano e quale dell’algoritmo.

L’obiettivo è di far pensare al **Discriminatore** che le nuove immagini sono opera di pittori umani.

# AlphaGo

AlphaGo (Lee) è un algoritmo di intelligenza artificiale basato su tecniche di apprendimento automatico sviluppato da **Google DeepMind**. È stato il primo software in grado di sconfiggere un maestro umano su un goban di dimensioni standard (19×19).



Il suo successore **AlphaGo Zero** **non usa conoscenze umane** ma si è addestrato da solo ed è **incredibilmente** molto più potente.

# AlphaZero

AlphaZero come AlphaGo Zero **non usa conoscenze umane**, ma **non è progettato per un gioco specifico** e (come nei fumetti di **Dragonball**) è ancora più potente.

Analogamente ai suoi predecessori impiega la **ricerca ad albero MonteCarlo** guidata da una **rete neurale convoluzionale profonda** addestrata per rinforzo.

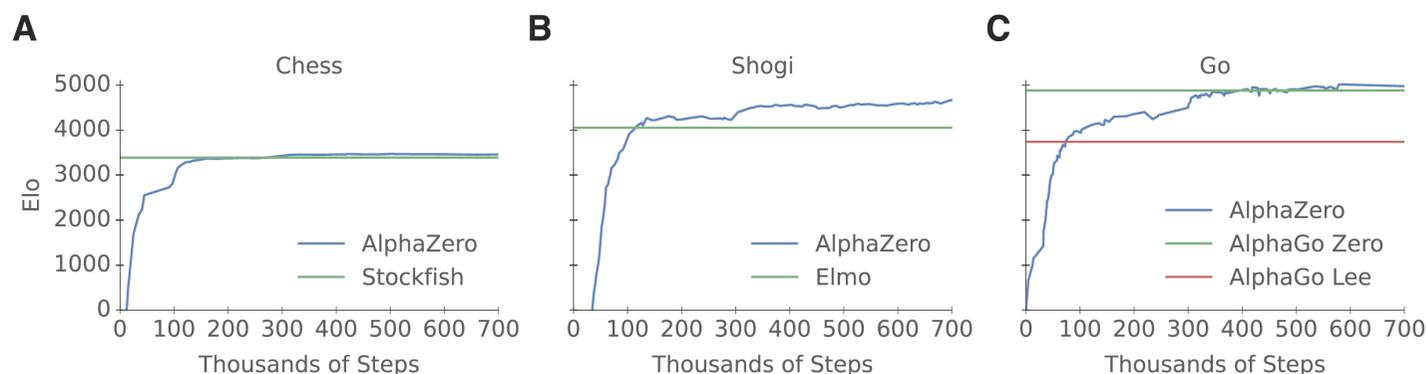
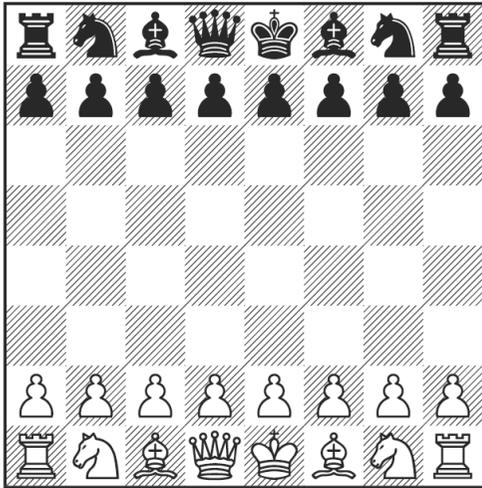


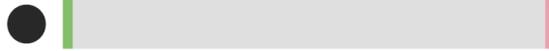
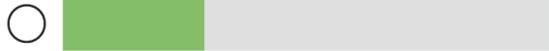
Figure 1: **Training AlphaZero for 700,000 steps.** Elo ratings were computed from games between different players where each player was given one second per move. **(A)** Performance of AlphaZero in chess, compared with the 2016 TCEC world-champion program Stockfish. **(B)** Performance of AlphaZero in shogi, compared with the 2017 CSA world-champion program Elmo. **(C)** Performance of AlphaZero in Go, compared with AlphaGo Lee and AlphaGo Zero (20 blocks over 3 days).

# Chess

## AlphaZero vs. Stockfish



W: 29.0% D: 70.6% L: 0.4%



W: 2.0% D: 97.2% L: 0.8%

# Shogi

## AlphaZero vs. Elmo



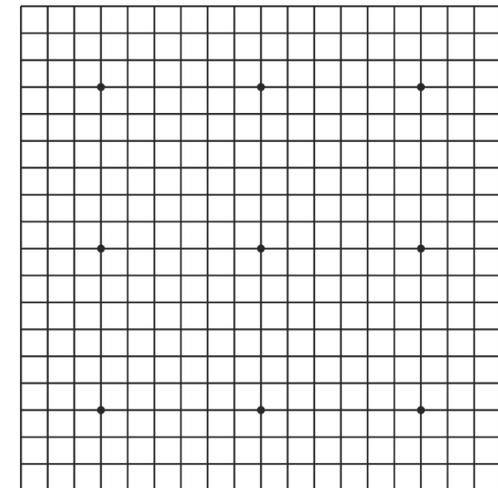
W: 84.2% D: 2.2% L: 13.6%



W: 98.2% D: 0.0% L: 1.8%

# Go

## AlphaZero vs. AG0



W: 68.9% L: 31.1%



W: 53.7% L: 46.3%

... in my observation, AlphaZero prioritizes piece activity over material, preferring positions that to my eye looked risky and aggressive. Programs usually reflect priorities and prejudices of programmers, but because AlphaZero programs itself, I would say that its style reflects the truth. This superior understanding allowed it to outclass the world's top traditional program despite calculating far fewer positions per second. It's the embodiment of the cliché, "work smarter, not harder." AlphaZero shows us that machines can be the experts, not merely expert tools. [Garry Kasparov, Science](#)

# Intelligenza globale?

- Per ora è presto ma non vi sono limiti teorici.
- Potenzialità = Non saper fare nulla
- La potenza del caso (**L'orologiaio cieco**)
- Non serve un computer fisico / potenza di calcolo (**CPU, GPU, TPU**)
- Evoluzione diviene **lamarckiana** (**trasmissione dei caratteri acquisiti**)

# È un bene o un male?

Nessuna ed entrambe le cose, **il progresso è inevitabile**, dal momento in cui è stata scheggiata la prima selce il destino dell'uomo era segnato (... il vaso di Pandora ...)

Ricordiamoci che (**purtroppo?**)

**Quello che è possibile è obbligatorio.**