Osservazioni geometriche per l'intelligenza artificiale

Patrizio Frosini

Dipartimento di Matematica - Università di Bologna
patrizio.frosini@unibo.it

Rome Centre on Mathematics for Modelling and Data ScienceS
20 marzo 2024

Sommario

Finalità di questo seminario

Il gioco di parole: *"Osservazioni geometriche"* può essere inteso in due modi:

- Considerazioni di tipo **geometrico**
- La geometria delle **osservazioni**

Entrambi questi significati sono di nostro interesse nella ricerca sull'intelligenza artificiale.

In questo seminario vorremmo mostrare come la geometria e la topologia rendano disponibili idee e strumenti che potrebbero essere utili per studiare lo **spazio degli osservatori** *nella ricerca sull'intelligenza artificiale.*

L'esposizione sarà, per quanto possibile, informale.

Finalità di questo seminario

Le idee principali:

- Lo spazio degli osservatori è spesso più importante dello spazio dei dati;
- Per studiare lo spazio degli osservatori è utile sviluppare un nuovo modello geometrico;
- Lo sviluppo di un nuovo modello geometrico offre dei vantaggi teorici e applicativi.

In altre parole cercherò di illustrare come si possa munire lo spazio degli osservatori di una struttura geometrica, in modo da poterlo studiare e usare in modo più unitario, efficiente e significativo. Si tratta, dunque, di interpretare il concetto di osservatore come una variabile nel trattamento e nell'interpretazione dei dati.

Cos'è l'intelligenza?

L'intelligenza può essere descritta come la capacità di percepire o inferire informazioni e di conservarle come conoscenza da applicare a comportamenti adattativi all'interno di un ambiente o contesto.

WIKIPEDIA

(Si veda anche *Sternberg, R. J. & Salter, W. (1982). Conceptions of intelligence. In Sternberg, R. J. (Ed.), Handbook of human intelligence.*)

Ma non tutti i comportamenti adattativi sono intelligenti...

Il ruolo degli osservatori nello studio dell'intelligenza

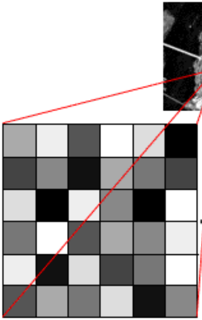
Ci interessano in particolare queste domande:

- Come è possibile formalizzare il ruolo dell'osservatore in ambito topologico-geometrico?
- Quali sono le principali proprietà topologico-geometriche dello spazio degli osservatori?
- Cosa vuol dire deformare un osservatore e il suo comportamento in relazione ai dati?
- Come è possibile muoversi nello spazio degli osservatori?
- Quali conseguenze può avere lo studio del concetto di osservatore nell'ambito del machine learning?

Sommario

I dati sono spesso rappresentabili come funzioni

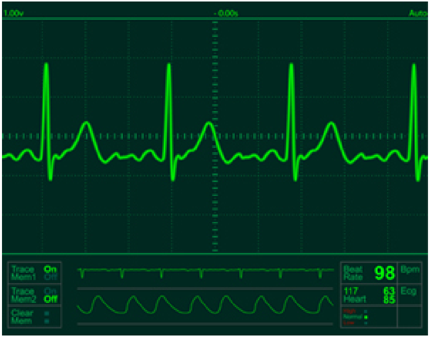
I dati sono spesso descritti da funzioni a valori reali o vettoriali, cioè leggi che associano agli oggetti dei valori numerici o delle *n*-uple di valori numerici.
Facciamo alcuni esempi:



170	238	85	255	221	0
68	136	17	170	119	68
221	0	238	136	0	255
119	255	85	170	136	238
238	17	221	68	119	255
85	170	119	221	17	136

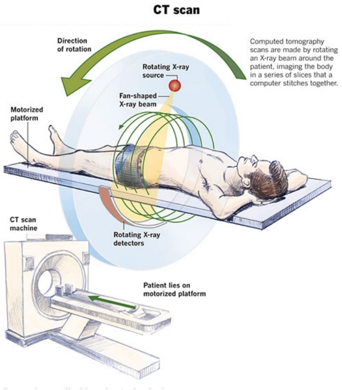
IMMAGINI

I dati sono spesso rappresentabili come funzioni



ELETTROCARDIOGRAMMI

I dati sono spesso rappresentabili come funzioni



TOMOGRAFIA ASSIALE COMPUTERIZZATA

I dati sono spesso rappresentabili come funzioni

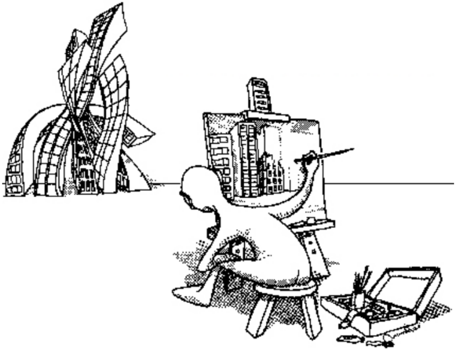
Nel nostro modello i dati sono descritti da funzioni a valori reali o vettoriali.

Indichiamo con Φ l'insieme dei dati ammissibili, cioè delle funzioni che possono essere interpretate come segnali da elaborare.

È importante osservare che solo alcune funzioni descrivono dati ammissibili: ad esempio, una funzione dal piano reale ai numeri reali che rappresenti un'immagine a livelli di grigio dovrà assumere valori compresi in un intervallo delimitato. Se questo non accade, la funzione non viene usualmente interpretata come la rappresentazione di un'immagine.

I dati vengono elaborati dagli osservatori

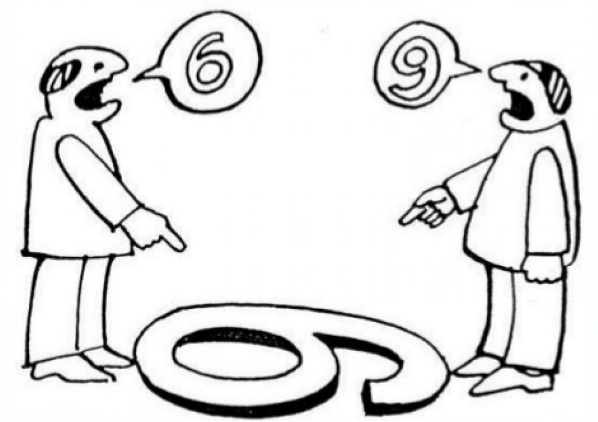
I dati hanno significato solo se c'è un osservatore che li elabora.



Un osservatore è un agente che trasforma dati in altri dati rispettandone le equivalenze e (usualmente) semplificandone la struttura.

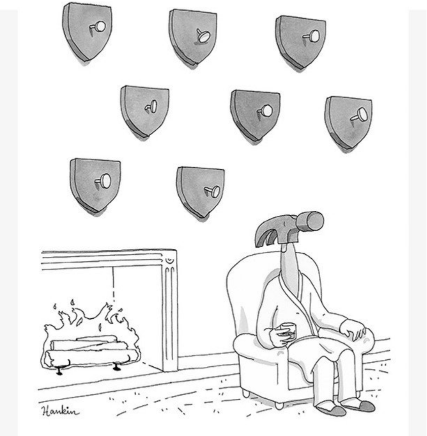
L'osservatore è una variabile del problema

L'interpretazione dei dati dipende fortemente dall'osservatore scelto.



Non c'è alcuna struttura nei dati

In generale, non c'è alcuna struttura nei dati. La struttura dei dati è una proiezione della struttura dell'osservatore.



La coppia (dato, osservatore)

In realtà il nostro interesse non è quasi mai rivolto direttamente ai dati, ma alla relazione fra i dati e l'osservatore. Quel che ci importa è soprattutto il modo in cui l'osservatore reagisce alla presenza di informazioni.

I casi in cui l'interesse per i dati sembra centrale sono quelli in cui la modalità di reazione degli osservatori è tacitamente condivisa.
Tutto ciò induce a privilegiare lo studio della "forma degli osservatori" allo studio della "forma dei dati".

Certe trasformazioni non sono rilevanti

Alcune trasformazioni sono irrilevanti per l'osservatore.

L'analisi di un globulo rosso non cambia se applichiamo una rotazione.

Un modello matematico per gli osservatori

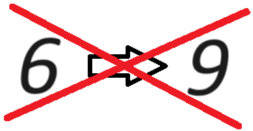
Gli osservatori possono essere visti come operatori che prendono in ingresso dati e li trasformano. In questo seminario parlerò di approssimazione di osservatori tramite operatori equivarianti non espansivi (in inglese **G**roup **E**quivariant **N**on-**E**xpansive **O**perators: **GENEO**).

Ci interessa particolarmente studiare la geometria degli spazi di osservatori rappresentati da GENEO.

Coppie di percezione

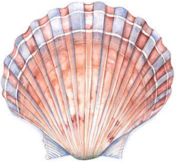
Chiameremo *coppia di percezione* ogni coppia (Φ, G) dove Φ sia uno spazio arbitrario di funzioni che rappresentano i dati ammissibili e G sia un gruppo di trasformazioni che preservano Φ.

In parole povere, vogliamo che ogni dato ammissibile venga trasformato da ogni trasformazione in G in un altro dato ammissibile. Per esempio, la traslazione temporale di un elettrocardiogramma è ancora un elettrocardiogramma e la rotazione nel piano della foto di una conchiglia è ancora l'immagine di una conchiglia. Invece la rotazione di 180 gradi dell'immagine di un 6 non è più l'immagine di un 6.



Il gruppo di equivarianza come variabile del modello

Il gruppo di equivarianza varia con l'osservatore:

forma di  **= forma di** 

forma di M $\neq$ **forma di** W

Operatori equivarianti fra coppie di percezione

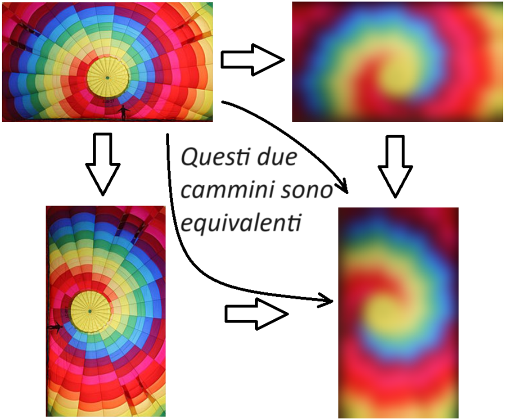
Un **operatore equivariante** *dalla coppia di percezione* (Φ, G) *alla coppia di percezione* (Ψ, K) *rispetto all'omomorfismo* $T : G \to K$ *è una legge che trasforma le funzioni appartenenti all'insieme di segnali ammissibili* Φ *in funzioni appartenenti all'insieme di segnali ammissibili* Ψ, *in modo tale da commutare con le trasformazioni appartenenti al gruppo. Ciò significa che applicare prima la trasformazione* g *e poi l'operatore* F *dà lo stesso risultato che applicare prima l'operatore* F *e poi la trasformazione* $T(g)$.

$$F(\varphi \circ g) = F(\varphi) \circ T(g)$$

Nel nostro contesto ogni osservatore è un operatore che agisce sui dati trasformandoli in modo equivariante.
Useremo l'acronimo GEO per indicatore ogni operatore equivariante (Group Equivariant Operator).

Operatori equivarianti fra coppie di percezione

ESEMPIO: L'operatore che sfuoca le immagini (usando un nucleo di convoluzione rotazionalmente simmetrico) è un operatore equivariante rispetto al gruppo delle rotazioni del piano intorno a un punto.

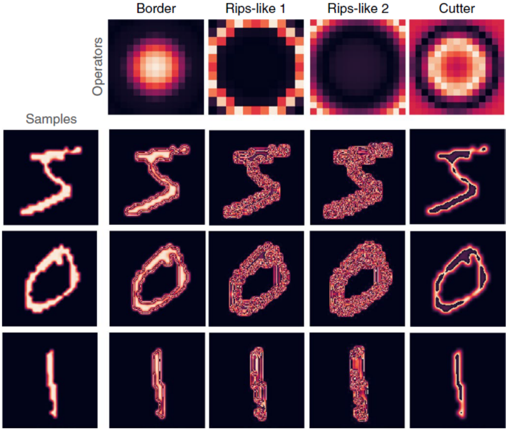


Operatori non espansivi

In molti casi siamo interessati a **operatori non espansivi**, cioè operatori che conservano o fanno diminuire la distanza fra i dati.
Il motivo è che gli operatori più utili sono quelli che, in un certo senso, semplificano le relazioni fra i dati "dimenticando" parte dell'informazione inizialmente disponibile.
Vogliamo quindi che la distanza fra i dati in uscita sia minore o uguale a quella dei dati in entrata.
Useremo l'acronimo **GENEO** per indicatore ogni operatore equivariante non espansivo (Group Equivariant Non-Expansive Operator).

Altri esempi di GENEO

Alcuni operatori equivarianti non espansivi su MNIST (modified National Institute of Standards and Technology database):



In sintesi

- I dati si possono spesso interpretare come **funzioni** (p.e., un'immagine a livelli di grigio può essere vista come una funzione dal piano reale nei numeri reali, dove ogni valore definisce il livello di grigio dell'immagine in un punto).
- Un osservatore può essere visto come un **operatore** che trasforma dati in altri dati.
- Gli osservatori interessanti hanno spesso proprietà di **equivarianza**, cioè commutano rispetto all'azione di un gruppo.
- Gli osservatori hanno spesso la proprietà di semplificare le relazioni fra i dati: questo fatto può essere parafrasato dicendo che non incrementano le distanze fra i dati. In questo caso si dicono **non espansivi**.

I GENEO nel machine learning

Il confronto di dati non è quasi mai un processo diretto: è quasi sempre mediato dal confronto fra dati prodotti da agenti/operatori tramite opportune trasformazioni dei dati originali.

- L'importanza degli operatori equivarianti nel machine learning è stata evidenziata da vari autori (Mallat, Poggio, Rosasco...)

Un elemento innovativo nella ricerca su questi operatori è costituito dallo sviluppo di tecniche topologiche e geometriche che tengano conto delle peculiarità informative espresse dall'**equivarianza** e dalla **non espansività**.

Non si tratta di studiare tali operatori come entità singole, ma di indagare le proprietà geometriche e topologiche degli spazi di tali operatori.

Due utili proprietà

Se gli spazi dei dati sono compatti, allora anche lo spazio degli osservatori è compatto.

QUINDI

Se gli spazi dei dati sono compatti (e dunque approssimabili con un errore arbitrariamente piccolo tramite un insieme finito), allora lo spazio degli osservatori è approssimabile con un errore arbitrariamente piccolo tramite un insieme finito di osservatori.

Due utili proprietà

Se lo spazio dei dati in uscita è convesso, allora anche lo spazio degli osservatori è convesso.

QUINDI

Se gli spazi dei dati sono compatti e convessi, allora anche lo spazio degli osservatori è compatto e convesso e dunque ogni funzione di costo strettamente convessa ammette un solo minimo sullo spazio degli osservatori.

Quindi lo spazio dei GENEO è particolarmente adatto ai processi di approssimazione e minimizzazione.

Approssimazione di osservatori

IL NOSTRO OBIETTIVO PRINCIPALE: APPROSSIMARE OSSERVATORI IDEALI

Nel nostro modello "approssimare un osservatore" significa cercare un GENEO F che minimizzi un'opportuna "funzione costo" $c(F)$.

La funzione costo quantifica l'errore che si commette prendendo il GENEO F al posto dell'osservatore ideale.

Dato che lo spazio dei GENEO è compatto e convesso (sotto l'ipotesi che gli spazi dei dati siano compatti e convessi), se la funzione costo $c(F)$ è strettamente convessa abbiamo che esiste uno e un solo GENEO che approssima al meglio l'osservatore ideale.

Un'osservazione importante

Mentre molto spesso non ha senso considerare le medie dei dati originali (perché l'insieme dei dati ammissibili è frequentemente non convesso), l'assunzione che il codominio dei GENEO sia convesso permette di considerare la media degli osservatori.

Costruire GENEO

frontiers in Artificial Intelligence

ORIGINAL RESEARCH
published: 15 February 2022
doi: 10.3389/frai.2022.786091

Check for updates

On the Construction of Group Equivariant Non-Expansive Operators *via* Permutants and Symmetric Functions

Francesco Conti [1,2], *Patrizio Frosini* [3,4,5,6*] *and Nicola Quercioli* [3,7]

[1] *Department of Mathematics, University of Pisa, Pisa, Italy,* [2] *Institute of Information Science and Technologies "A. Faedo", National Research Council of Italy (CNR), Pisa, Italy,* [3] *Department of Mathematics, University of Bologna, Bologna, Italy,* [4] *Alma Mater Research Center on Applied Mathematics, University of Bologna, Bologna, Italy,* [5] *Alma Mater Research Institute for Human-Centered Artificial Intelligence, University of Bologna, Bologna, Italy,* [6] *Research Centre on Electronic Systems for the Information and Communication Technology, University of Bologna, Bologna, Italy,* [7] *ENEA Centro Ricerche Bologna, Bologna, Italy*

https://www.frontiersin.org/articles/10.3389/frai.2022.786091/full

Costruzione di GENEO con metodi elementari

Proposizione (Composizione)

Se $F_1 \in \mathrm{GENEO}((\Phi, G_\Phi), (\Psi, G_\Psi))$ *rispetto a* $T_1 : G_\Phi \to G_\Psi$ *e* $F_2 \in \mathrm{GENEO}\big((\Psi, G_\Psi), (\chi, G_\chi)\big)$ *rispetto a* $T_2 : G_\Psi \to G_\chi$ *allora* $F_2 \circ F_1 \in \mathrm{GENEO}\big((\Phi, G_\Phi), (\chi, G_\chi)\big)$ *rispetto a* $T_2 \circ T_1 : G_\Phi \to G_\chi$.

Proposizione (Immagine tramite una funzione 1-Lipschitz)

Se $F_1, \ldots, F_n \in \mathrm{GENEO}((\Phi, G_\Phi), (\Psi, G_\Psi))$ *rispetto a* $T : G_\Phi \to G_\Psi$, L *è una mappa 1-Lipschitz da* $\mathbb{R}^n$ *a* $\mathbb{R}$, *e* $L^*(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$ *(dove* L^* *è la mappa indotta da* L*), allora* $L^*(F_1, \ldots, F_n) \in \mathrm{GENEO}((\Phi, G_\Phi), (\Psi, G_\Psi))$ *rispetto a* T.

Da quest'ultima proposizione seguono le tre affermazioni seguenti.

Costruzione di GENEO con metodi elementari

Proposizione (RETICOLI DI GENEO)

Se $F_1, \ldots, F_n \in \mathrm{GENEO}((\Phi, G_\Phi), (\Psi, G_\Psi))$ *r. a* $T : G_\Phi \to G_\Psi$ *e* $\max(F_1, \ldots, F_n)(\Phi), \min(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$, *allora* $\max(F_1, \ldots, F_n), \min(F_1, \ldots, F_n) \in \mathrm{GENEO}((\Phi, G_\Phi), (\Psi, G_\Psi))$ *r. a* T.

Proposizione (Traslazione)

Se $F \in \mathrm{GENEO}((\Phi, G_\Phi), (\Psi, G_\Psi))$ *r. a* $T : G_\Phi \to G_\Psi$, *e* $F_b(\Phi) \subseteq \Psi$ *per* $F_b(\varphi) := F(\varphi) - b$, *allora* $F_b \in \mathrm{GENEO}((\Phi, G_\Phi), (\Psi, G_\Psi))$ *r. a* T.

Proposizione (Combinazione convessa)

Se $F_1, \ldots, F_n \in \mathrm{GENEO}((\Phi, G_\Phi), (\Psi, G_\Psi))$ *r. a* $T : G_\Phi \to G_\Psi$, $(a_1, \ldots, a_n) \in \mathbb{R}^n$ *con* $\sum_{i=1}^n |a_i| \leq 1$ *e* $F_\Sigma(\Phi) \subseteq \Psi$ *per* $F_\Sigma(\varphi) := \sum_{i=1}^n a_i F_i(\varphi)$, *allora* $F_\Sigma \in \mathrm{GENEO}((\Phi, G_\Phi), (\Psi, G_\Psi))$ *r. a* T.

Misure permutanti

Consideriamo l'insieme $\Phi = \mathbb{R}^X \cong \mathbb{R}^n$ di tutte le funzioni da un insieme finito $X = \{x_1, \ldots, x_n\}$ a $\mathbb{R}$, e un sottogruppo G del gruppo $\mathrm{Bij}(X)$ di tutte le permutazioni di X.

Definizione

Una misura finita (con segno) μ su $\mathrm{Bij}(X)$ è detta una misura permutante rispetto a G se ogni sottoinsieme H di $\mathrm{Bij}(X)$ è misurabile e μ è invariante sotto l'azione di coniugio di G (cioè, $\mu(H) = \mu(gHg^{-1})$ per ogni $g \in G$).

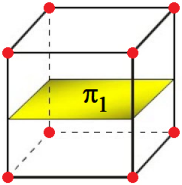
Proposizione

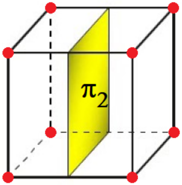
Se μ è una misura permutante rispetto a G, allora la mappa $F_\mu : \mathbb{R}^X \to \mathbb{R}^X$ definita ponendo $F_\mu(\varphi) := \sum_{h \in \mathrm{Bij}(X)} \varphi h^{-1} \, \mu(h)$ è un GEO lineare. Se $\sum_{h \in \mathrm{Bij}(X)} |\mu(h)| \leq 1$, allora $F_\mu(\varphi)$ è un GENEO.

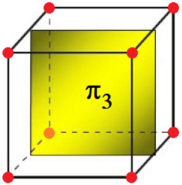
Un esempio di misura permutante

Consideriamo l'insieme X dei vertici di un cubo in $\mathbb{R}^3$, e il gruppo G delle isometrie di $\mathbb{R}^3$ che portano X in X e conservano l'orientazione. Siano π_1, π_2, π_3 i tre piani che contengono il baricentro di X e sono paralleli a una faccia del cubo. Sia $h_i : X \to X$ la simmetria ortogonale rispetto a π_i, per $i \in \{1, 2, 3\}$.

Possiamo ora definire una misura permutante μ sul gruppo $\mathrm{Bij}(X)$ ponendo $\mu(h_1) = \mu(h_2) = \mu(h_3) = c$, dove c è un numero reale positivo, e $\mu(h) = 0$ per ogni $h \in \mathrm{Bij}(X)$ con $h \notin \{h_1, h_2, h_3\}$.







GENEO lineari e misure permutanti

Annals of Mathematics and Artificial Intelligence
https://doi.org/10.1007/s10472-022-09830-1

On the finite representation of linear group equivariant operators via permutant measures

Giovanni Bocchi[1] · Stefano Botteghi[2] · Martina Brasini[2] · Patrizio Frosini[2] · Nicola Quercioli[3]

Accepted: 26 December 2022

https://rdcu.be/c5Obw

Teorema di rappresentazione per i GENEO lineari

Questo teorema rafforza la proposizione precedentemente vista sulla costruzione di GENEO lineari tramite misure permutanti.

Teorema

Assumiamo che $G \subseteq \mathrm{Bij}(X)$ *agisca transitivamente sull'insieme finito* X *e che* F *sia una mappa da* $\mathbb{R}^X$ *a* $\mathbb{R}^X$. *La mappa* F *è un GENEO lineare da* $\mathbb{R}^X$ *a* $\mathbb{R}^X$ *rispetto all'omomorfismo identico* Id_G **se e solo se** *esiste una misura permutante* μ *rispetto a* G, *tale che* $F(\varphi) = \sum_{h \in \mathrm{Bij}(X)} \varphi h^{-1} \mu(h)$ *per ogni* $\varphi \in \mathbb{R}^X$, *e* $\sum_{h \in \mathrm{Bij}(X)} |\mu(h)| \leq 1$.

NB: L'insieme $\mathrm{PM}(G)$ delle misure permutanti rispetto a G è un reticolo e uno spazio vettoriale reale.

Perché costruire GENEO tramite misure permutanti?

Osserviamo che più il supporto di una misura permutante μ è piccolo, più la sommatoria $F_\mu(\varphi) := \sum_{h \in \mathrm{Bij}(X)} \varphi h^{-1} \ \mu(h)$ che definisce il GENEO associato è semplice da calcolare. Nell'esempio appena visto il gruppo $\mathrm{Bij}(X)$ ha cardinalità 40.320, il gruppo di equivarianza G contiene 24 elementi, mentre il supporto della misura permutante contiene soltanto 3 permutazioni di X.

L'utilità del metodo costruttivo basato sulle misure permutanti sta nel fatto che spesso possiamo usare permutanti piuttosto piccoli.

Costruzione di GENEO non lineari

Il metodo per costruire GENEO basati su misure permutanti può essere generalizzato sostituendo la media aritmetica con un'altra funzione simmetrica. Per farlo, ricordiamo il concetto di *permutante*, che equivale a quello di *misura permutante uniformemente distribuita sul suo supporto*.

Definizione

Un insieme $H \subseteq \mathrm{Bij}(X)$ *è un* **permutante** *per* G *se* $H = \emptyset$ *o* $gHg^{-1} = H$ *per ogni* $g \in G$.

Nel prossimo lucido vedremo che quando abbiamo una funzione simmetrica e un **permutante** per il gruppo di equivarianza G, possiamo facilmente costruire un GENEO (generalmente non lineare) da (Φ, G) a $(\mathbb{R}^X, G)$ rispetto all'omomorfismo identico $\mathrm{id}_G : G \to G$.

Costruzione di GENEO non lineari

Sia $\mathscr{S}: \mathbb{R}^n \to \mathbb{R}$ una funzione simmetrica. Se $H = \{h_i\}_{i=1}^n$ è un permutante non vuoto per $G \subseteq \mathrm{Bij}_\Phi(X)$, allora possiamo definire un operatore $\mathscr{S}_H : \Phi \to \mathbb{R}^X$ ponendo, per ogni $\varphi \in \Phi$,

$$\mathscr{S}_H(\varphi) := \mathscr{S}(\varphi \circ h_1, \ldots, \varphi \circ h_n),$$

dove $\mathscr{S}(\varphi \circ h_1, \ldots, \varphi \circ h_n)(x) := \mathscr{S}((\varphi \circ h_1)(x), \ldots, (\varphi \circ h_n)(x))$ per ogni $x \in X$.

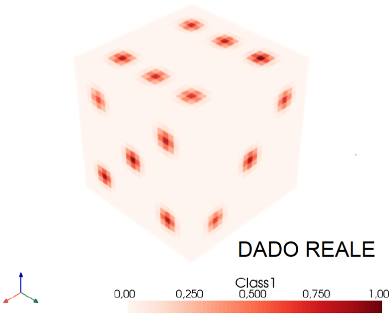
Proposizione

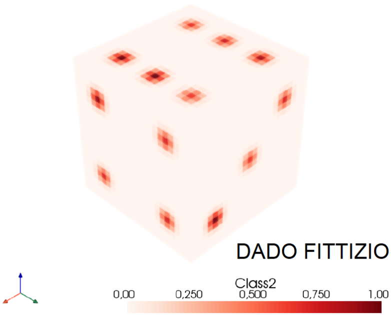
$\mathscr{S}_H$ è un GEO da (Φ, G) a $(\mathbb{R}^X, G)$ rispetto all'omomorfismo identico $\mathrm{id}_G : G \to G$. Se la restrizione di $\mathscr{S}$ a $\mathrm{Im}(\Phi)^n$ è non espansiva, allora $\mathscr{S}_H$ è un GENEO da (Φ, G) a $(\mathbb{R}^X, G)$ rispetto a id_G.

Sommario

Cosa succede quando applichiamo GENEO ai dati?

Un esempio di utilizzo: confronto tra dadi veri e dadi falsi.

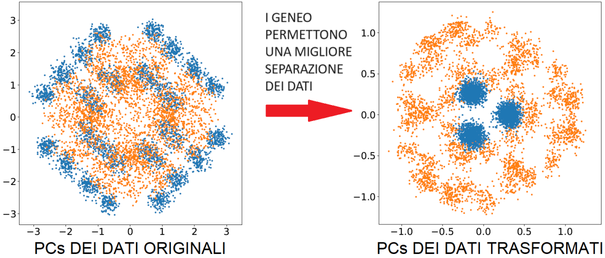




(Esperimento e calcoli di Giovanni Bocchi)

Cosa succede quando applichiamo GENEO ai dati?

Abbiamo prodotto 10000 dadi (un training set di cardinalità 7000 e un test set di cardinalità 3000), abbiamo poi applicato la PCA al test set e al test set trasformato da un GENEO ottimizzato sul training set:



Per ogni matrice vengono riportate le prime due componenti principali. I punti blu sono associati ai **dadi reali**, mentre quelli arancioni a **dadi falsi**. Il GENEO che utilizziamo è una combinazione convessa di 3 GENEO definiti da misure permutanti.

Un'applicazione reale: ricerca di tasche nelle proteine

GENEOnet: A new machine learning paradigm based on Group Equivariant Non-Expansive Operators. An application to protein pocket detection.

Giovanni Bocchi [1], Patrizio Frosini [2], Alessandra Micheletti [1], Alessandro Pedretti [3] Carmen Gratteri [4], Filippo Lunghini [5] , Andrea Rosario Beccari [5] and Carmine Talarico [5]

[1] Department of Environmental Science and Policy, Università degli Studi di Milano

[2] Department of Mathematics, Università degli Studi di Bologna

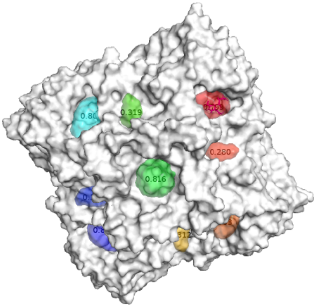
[3] Department of Pharmaceutical Sciences, Università degli Studi di Milano

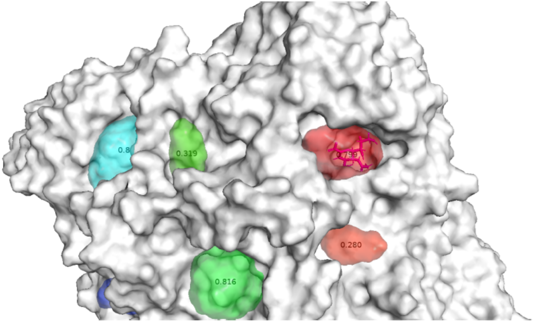
[4] Dipartimento di Scienze della Salute, Università degli Studi "Magna Græcia di Catanzaro"

[5] Dompé Farmaceutici SpA

https://arxiv.org/ftp/arxiv/papers/2202/2202.00451.pdf

Un’applicazione reale: ricerca di tasche nelle proteine

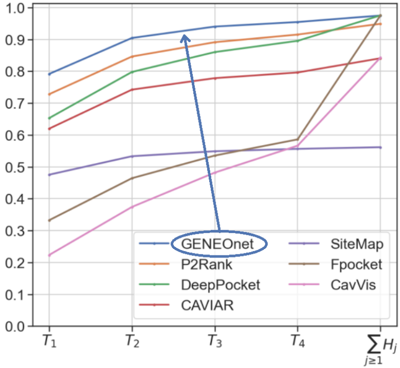




La ricerca delle tasche nelle proteine può essere effettuata identificando un GENEO ottimale nell’inviluppo convesso di 8 GENEO (ciascuno focalizzato su una particolare proprietà delle tasche).

Un'applicazione reale: ricerca di tasche nelle proteine

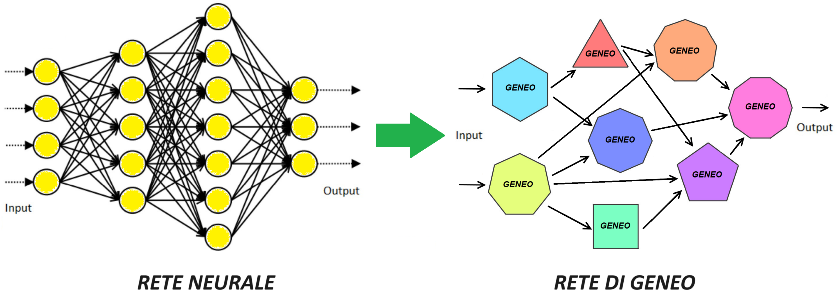
Ecco i risultati dei nostri esperimenti:



Occorre tenere presente che GENEOnet utilizza 17 parametri, mentre una CNN come DeepPocket richiede 665122 parametri.

Uso dei GENEO nel Machine Learning

In prospettiva, stiamo cercando di sviluppare una buona teoria composizionale per costruire reti efficienti di GENEO. Alcuni esperimenti preliminari suggeriscono che la sostituzione dei neuroni con GENEO potrebbero rendere il deep learning più trasparente e interpretabile e velocizzare il processo di apprendimento.



Sommario

Intelligenza e contraddizione

La rappresentazione operatoriale
del concetto di osservatore
ha un'altra importante conseguenza:
una sorta di «principio di contraddizione».

ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect

Cognitive Systems Research 10 (2009) 297–315

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Does intelligence imply contradiction?

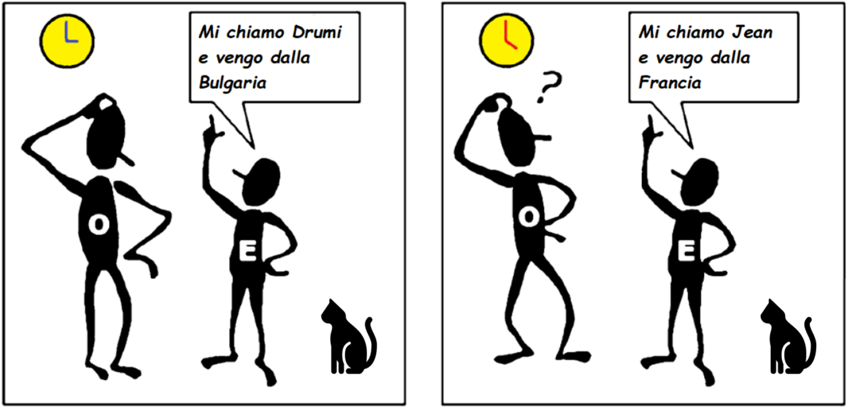
Action editor: Vasant Honavar

P. Frosini *

Department of Mathematics and ARCES, University of Bologna, I-40126 Bologna, Italy

Intelligenza e contraddizione

Cosa intendiamo per contraddizione?



Intelligenza e contraddizione

Teorema: Ogni entità sufficientemente intelligente è contraddittoria.

Ludwig Josef Johann Wittgenstein (Vienna, 26 aprile 1889 – Cambridge, 29 aprile 1951)

Tractatus
Logico-Philosophicus

By
LUDWIG WITTGENSTEIN

With an Introduction by
BERTRAND RUSSELL, F.R.S.

NEW YORK
HARCOURT, BRACE & COMPANY, INC.
LONDON: KEGAN PAUL, TRENCH, TRUBNER & CO., LTD.
1922

Le sette asserzioni principali

1. Il mondo è tutto ciò che accade.
2. Ciò che accade, il fatto, è il sussistere di stati di cose.
3. L'immagine logica dei fatti è il pensiero.
4. Il pensiero è la proposizione munita di senso.
5. La proposizione è una funzione di verità delle proposizioni elementari.
6. La forma generale della funzione di verità è: $[\bar{p}, \bar{\xi}, N(\bar{\xi})]$. Questa è la forma generale della proposizione.
7. Su ciò di cui non si può parlare, si deve tacere.

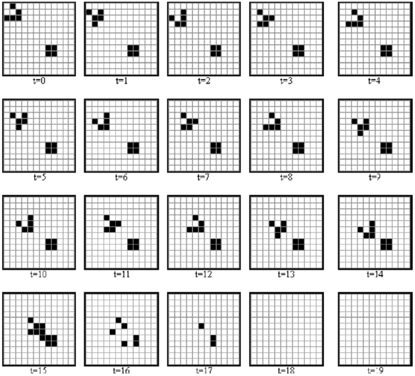
Intelligenza e contraddizione

In modo equivalente possiamo dire che

> Il comportamento di ogni entità sufficientemente intelligente è impredicibile.

Intelligenza e contraddizione

Come dimostrarlo?
È possibile un approccio basato sugli automi cellulari.



Intelligenza e contraddizione

Linea dimostrativa:

- Si individua un osservatore (inteso come un operatore che trasforma le funzioni che rappresentano gli stati dell'automa cellulare in funzioni che descrivono l'entità percepita e l'ambiente che la circonda).
- Si definisce l'intelligenza di una entità come la sua capacità di sopravvivere nell'ambiente secondo il giudizio dell'osservatore.
- Si prova che esiste una soglia per l'intelligenza (dipendente dal numero di stati che l'osservatore può associare all'entità e all'ambiente), oltre la quale l'entità osservata appare necessariamente contraddittoria all'osservatore prescelto.

In questo modello, quindi, la contraddittorietà e la non predicibilità non appaiono come limitazioni delle strutture intelligenti ma come condizioni necessarie per lo sviluppo di comportamenti intellettivi complessi.

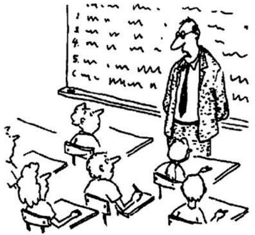
Intelligenza e contraddizione

Teorema. Sia E un'entità avente lunghezza finita di vita e si supponga che il suo ambiente sia deterministico. Se l'intelligenza di E è maggiore del prodotto delle cardinalità degli insiemi P_{ent} e P_{ENV} l'entità E deve essere necessariamente contraddittoria.

ATTENZIONE!: Il teorema non asserisce che le entità intelligenti devono cambiare il loro comportamento (fatto ovvio) ma che devono farlo senza che l'osservatore ne comprenda il motivo.

P_{ent} = insieme degli stati di E riconosciuti dall'osservatore.

P_{ENV} = insieme degli stati dell'ambiente riconosciuti dall'osservatore.



Mi aspetto che voi siate autonomi, creativi, critici e che facciate tutto quello che dico io.

Intelligenza e contraddizione

La formalizzazione matematica di questo approccio può essere trovata nell'articolo

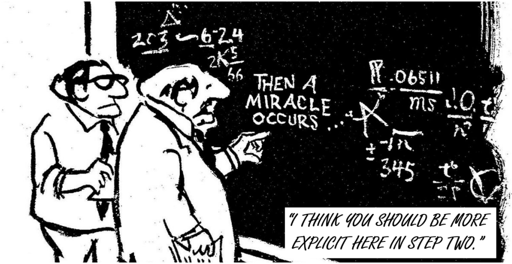
P. Frosini, Does intelligence imply contradiction?, Cognitive Systems Research, vol. 10 (2009), n. 4, 297-315.

(Chi lo desideri può trovare una presentazione dell'articolo realizzata da Mattia G. Bergomi al link *https://mgbergomi.github.io/Contradiction/* e un'introduzione informale all'argomento al link *http://maddmaths.simai.eu/divulgazione/langolo-arguto/lintelligenza-della-contraddizione/*.)

Il preprint dell'articolo è disponibile a questo link: *https://arxiv.org/pdf/0801.0232.pdf*.

Intelligenza e contraddizione

Abbiamo visto che ogni agente A appare imprevedibile agli occhi di un osservatore prefissato se l' *"intelligenza"* di A supera una soglia espressa dal prodotto del numero di stati che l'osservatore può percepire nell'agente e nel contesto ambientale. Ciò implica che per avere prevedibilità di comportamento occorre scegliere modelli in cui la predetta soglia sia più grande del valore di intelligenza desiderato.



Sommario

Il framework matematico

nature machine intelligence

ARTICLES

https://doi.org/10.1038/s42256-019-0087-3

Towards a topological–geometrical theory of group equivariant non-expansive operators for data analysis and machine learning

Mattia G. Bergomi[1], Patrizio Frosini[2,3*], Daniela Giorgi[4] and Nicola Quercioli[2,3]

We provide a general mathematical framework for group and set equivariance in machine learning. We define group equivariant non-expansive operators (GENEOs) as maps between function spaces associated with groups of transformations. We study the topological and metric properties of the space of GENEOs to evaluate their approximating power and set the basis for general strategies to initialize and compose operators. We define suitable pseudo-metrics for the function spaces, the equivariance groups and the set of non-expansive operators. We prove that, under suitable assumptions, the space of GENEOs is compact and convex. These results provide fundamental guarantees in a machine learning perspective. By considering isometry-equivariant non-expansive operators, we describe a simple strategy to select and sample operators. Thereafter, we show how selected and sampled operators can be used both to perform classical metric learning and to inject knowledge in artificial neural networks.

https://rdcu.be/bP6HV

Uso dei GENEO nel Machine Learning

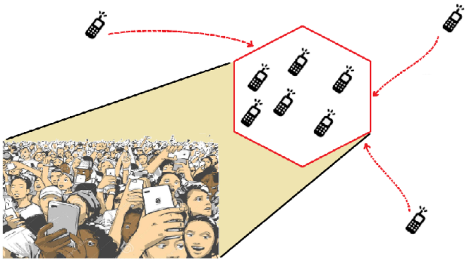
Per maggiori dettagli sull'uso di GENEO nel Machine Learning:

- A. Micheletti, *A new paradigm for artificial intelligence based on group equivariant non-expansive operators*, In: *EMS Magazine*, 128 (2023), pp. 4–12.
- https://ems.press/content/serial-article-files/27673

Alcuni progetti ricerca

CNIT / WiLab - Huawei Joint Innovation Center (JIC)

Project on GENEOs for 6G



WILAB HUAWEI

Alcuni progetti ricerca

Horizon Europe (HORIZON)
Call: HORIZON-CL4-2023-HUMAN-01-CNECT
Project: 101135775 — PANDORA
Funding: approximately 9 million euros.

Task 3.3 - Leveraging domain knowledge for explainable learning: This task aims to investigate the use of domain knowledge in the development of explainable AI models. Tools like GENEOs for applications in TDA and ML and new theoretical methods of GENEOs for explainable AI will be used.

Sommario

Conclusioni

- Nella ricerca sull'apprendimento automatico potrebbe essere interessante spostare l'attenzione dallo spazio dei dati allo spazio degli osservatori.
- Questo cambiamento di prospettiva richiede lo sviluppo di nuovi modelli topologici e geometrici adatti a studiare la "forma" degli spazi di osservatori.
- L'approccio che abbiamo descritto ci offre vari risultati potenzialmente utili nella ricerca sull'intelligenza artificiale.

GRAZIE PER L'ATTENZIONE